

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification 6 : C12N 15/31, 15/00, 15/11, C07K 14/315, C07H 21/04, A61K 39/40, 39/395, G01N 33/569, G11B 5/00</p>	<p>A1</p>	<p>(11) International Publication Number: <b>WO 98/26072</b>  (43) International Publication Date: 18 June 1998 (18.06.98)</p>
<p>(21) International Application Number: PCT/US97/22578 (22) International Filing Date: 9 December 1997 (09.12.97) (30) Priority Data: 60/036,281 13 December 1996 (13.12.96) US (71) Applicant (for all designated States except US): ELI LILLY AND COMPANY [US/US]; Lilly Corporate Center, Indi- anapolis, IN 46285 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): BALTZ, Richard, H. [US/US]; 6438 North Olney, Indianapolis, IN 46220 (US). BURGETT, Stanley, G. [US/US]; 818 Bitter Bark Lane, In- dianapolis, IN 46227 (US). DEHOFF, Bradley, S. [US/US]; 7727 Maradona Drive South, Indianapolis, IN 46214 (US). HOSKINS, Jo, Ann [US/US]; 8229 Tern Court, Indianapo- lis, IN 46256 (US). JASKUNAS, Stanley, R., Jr. [US/US]; 16 Algonquian Drive, Natick, MA 07160 (US). MILLS, Bradley, J. [US/US]; 321 South Poplar, Morristown, IN 46161 (US). NORRIS, Franklin, H. [US/US]; 5518 South Arlington Avenue, Indianapolis, IN 46237 (US). PEERY, Robert, B. [US/US]; 372 Sycamore Street, Brownsburg, IN 46112 (US). ROCKEY, Pamela, K. [US/US]; 679 Ca-</p>		<p>nary Creek Drive, Franklin, IN 46131 (US). ROSTECK, Paul, R., Jr. [US/US]; 4247 Saffron Drive, Indianapolis, IN 46237 (US). SKATRUD, Paul, L. [US/US]; 5233 Tony Drive, Greenwood, IN 46143 (US). SMITH, Michele, C. [US/US]; 4515 North Lakeridge Drive, Indianapolis, IN 46234 (US). SOLENBERG, Patricia, J. [US/US]; 10116 Seabreeze Way, Indianapolis, IN 46256 (US). TREAD- WAY, Patti, J. [US/US]; 4268 West Fox Ridge Avenue, Greenwood, IN 46143 (US). YOUNG BELLIDO, Michele, L. [CA/US]; 4925 Common Vista Court, Indianapolis, IN 46220 (US).  (74) Agents: WEBSTER, Thomas, D. et al.; Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285 (US).  (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>
<p>(54) Title: STREPTOCOCCUS PNEUMONIAE DNA SEQUENCES (57) Abstract  Nucleotide fragments of the genome of the bacterium <i>Streptococcus pneumoniae</i> are provided. Also provided are ORFs encoded by said genome.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

-1-

## STREPTOCOCCUS PNEUMONIAE DNA SEQUENCES

5        This invention provides DNA sequences from the  
      *Streptococcus pneumoniae* genome, and methods of use of DNA  
      fragments originating therefrom in a variety of biological  
      and pharmaceutical applications.

      The recent emergence of widespread antibiotic  
10    resistance in common pathogenic bacterial species has  
      justifiably alarmed the medical and research communities.  
      Frequently these organisms are co-resistant to several  
      different antibacterial agents. Particularly problematic has  
      been the emergence and rapid spread of penicillin resistance  
15    in *Streptococcus pneumoniae*, which frequently causes upper  
      respiratory tract infections. Resistance to penicillin in  
      this organism can be due to modifications of one or more of  
      the penicillin-binding proteins (PBPs). Combating the  
      phenomenon of increasing resistance to antibiotic agents  
20    among pathogenic organisms such as *Streptococcus pneumoniae*  
      will require intensified research into the fundamental  
      molecular biology of such organisms. Greater knowledge about  
      the molecular biology of pathogenic organisms will lead to  
      new antibacterial agents having novel and effective actions.

25        While inroads in the development of new antibiotics and  
      new targets for antibiotic compounds have been made with a  
      variety of microorganisms, progress has been less apparent  
      in *Streptococcus pneumoniae*. In part, *Streptococcus*  
      *pneumoniae* presents a special case because this organism is  
30    highly recombinogenic and readily takes up exogenous DNA  
      from its surroundings. Thus, the need for new antibacterial  
      compounds and new targets for antibacterial therapy in  
      *Streptococcus pneumoniae* is more acute than in other  
      organisms.

-2-

The present invention relates to the genome of *S. pneumoniae*. The genomic information disclosed by the present invention enables: (1) preparation of molecular  
5 hybridization probes for use in PCR amplification of genes and regulatory regions, physical mapping, sequencing, mutagenesis, and mutation analysis, (2) homology comparisons with the genomes and open reading frames (ORFs) of other  
10 *S. pneumoniae* wherein the mutation is targeted to any site or sites in the DNA sequence disclosed herein, (4) identification of *S. pneumoniae* promoters and other gene regulatory sequences, (5) identification of proteins/ORFs encoded by *S. pneumoniae*, (6) identification of virulence  
15 genes in *S. pneumoniae*, (7) determination of the biological function of proteins/ORFs and RNAs encoded by *S. pneumoniae*, (8) production of kits useful for determining gene function in the cell, and kits for isolating and analyzing genes that are mutated in antibiotic resistant clinical isolates of *S.*  
20 *pneumoniae*, (9) production of proteins and RNAs encoded by *S. pneumoniae*, (10) production of antibodies against proteins and other antigens encoded by *S. pneumoniae*, (11) methods to identify compounds that bind to proteins and RNAs encoded by *S. pneumoniae* as potential new antibiotic  
25 compounds.

In another embodiment the invention relates to substantially purified proteins encoded by the *S. pneumoniae* genome.

30 Table 1 summarizes the proteins and nucleic acids disclosed herein, contigs, SEQ ID NO's and predicted functions.



-3-

"Genome" refers to the full complement of chromosomal and extra-chromosomal DNA within a cell. The genome comprises the genetic blueprint for all proteins and RNAs encoded by the cell or organism.

5 "ORF" (i.e. "open reading frame") designates a region of genomic DNA beginning with a Met or other initiation codon and terminating with a translation stop codon, potentially encoding a protein product. "Partial ORF" means  
10 a portion of an ORF as disclosed herein such that the initiation codon, the stop codon, or both are not disclosed.

"DNA chip" or "Bio Chip" or "Bio DNA chip" refers to a solid matrix or support onto which is applied an array of oligonucleotides, or nucleotide sequences, or gene  
15 fragments, or genomic fragments, of *S. pneumoniae* which may further comprise a layer of *S. pneumoniae* cells suspended thereover in a semisolid medium such as agar or agarose.

"Consensus sequence" refers to an amino acid or nucleotide sequence that may suggest the biological function of a protein, DNA, or RNA molecule. Consensus sequences are  
20 identified by comparing proteins, RNAs, and gene homologs from different species.

"Contiguous fragment building" or "Contiguous fragment" or "Contig" refers to the process and result, respectively, by which a fragment of DNA is assembled from smaller  
25 constituent DNA fragments by arranging the constituent pieces in their correct order and register such that the resulting contiguous fragment accurately depicts the native DNA sequence from which the smaller fragments originated.

"Computer readable medium" includes, for example, a  
30 floppy disc, hard disc, random access memory, read only memory, and CD-ROM.

The terms "cleavage" or "restriction" of DNA refers to the catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in

-4-

the DNA (viz. sequence-specific endonucleases). The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors, and other requirements are used in the manner well known to one of ordinary skill in the art. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer or can be found in the literature.

"Diagnostics" as used herein relates to *in vitro* or *in vivo* diagnosis for disease states or biological status in mammals, preferably humans.

"Therapeutics" and "therapeutic/diagnostic combinations" means the treatment, or diagnosis and treatment, of disease states or biological status by *in vivo* administration to mammals, preferably humans, of compositions of the present invention, for example, antibodies.

"Essential genes" or "essential ORFs" or "essential proteins" refer to genomic information or the protein(s) or RNAs encoded therefrom, which, when disrupted by knockout mutation, or by other mutation, produce inviability in cells harboring said mutation.

"Non-essential genes" or "non-essential ORFs" or "non-essential proteins" refer to genomic information or the protein(s) or RNAs encoded therefrom, which, when disrupted by knockout mutation, or other mutation, do not result in inviability of cells harboring said mutation.

"Minimal gene set" refers to a genus of about 256 genes that are conserved among different bacteria such as *M. genitalium* and *H. influenzae*. The minimal gene set appears to be necessary and sufficient to sustain life. See e.g. A. Mushegian and E. Koonin, "A minimal gene set for cellular life derived by comparison of complete bacterial genomes" *Proc. Nat. Acad. Sci.* 93, 10268 - 273 (1996).

-5-

The term "fragment thereof" denotes a fragment of a nucleic acid molecule described herein, wherein said fragment comprises a region of contiguity within said nucleic acid of at least 15 base pairs. The term may also  
5 refer to a peptide of at least 5 contiguous amino acid residues of a protein disclosed herein.

The term "plasmid" refers to an extrachromosomal genetic element. The starting plasmids herein are either commercially available, publicly available on an  
10 unrestricted basis, or can be constructed from available plasmids in accordance with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

15 "Recombinant DNA cloning vector" as used herein refers to any autonomously replicating agent, including, but not limited to, plasmids and phages, comprising a DNA molecule to which one or more additional DNA segments can or have been added.

20 The term "recombinant DNA expression vector" as used herein refers to any recombinant DNA cloning vector, for example a plasmid or phage, in which a promoter and other regulatory elements are present to enable transcription of the inserted DNA.

25 The term "vector" as used herein refers to a nucleic acid compound used for introducing exogenous DNA into host cells. A vector comprises a nucleotide sequence which may encode one or more protein molecules. Plasmids, cosmids, viruses, and bacteriophages, in the natural state  
30 or which have undergone recombinant engineering, are examples of commonly used vectors.

The terms "complementary" or "complementarity" as used herein refers to the capacity of purine and pyrimidine nucleotides to associate through hydrogen bonding in double

-6-

stranded nucleic acid molecules. The following base pairs are complementary: guanine and cytosine; adenine and thymine; and adenine and uracil.

"Oligonucleotide" refers to a short polymeric  
5 nucleotide chain comprising from about 2 to 25 nucleotides.

"Isolated nucleic acid compound" refers to any RNA or DNA sequence, however constructed or synthesized, which is locationally distinct from its natural location.

A "primer" is a nucleic acid fragment which  
10 functions as an initiating substrate for enzymatic or synthetic elongation of a nucleic acid molecule.

The term "promoter" refers to a DNA sequence which directs transcription of DNA to RNA.

A "probe" as used herein is a labeled nucleic acid  
15 compound which can be used to hybridize with another nucleic acid compound.

The term "hybridization" or "hybridize" as used herein refers to the process by which a single-stranded nucleic acid molecule joins with a complementary strand  
20 through nucleotide base pairing.

"Recorded" as used herein refers to a process for storing information on a computer readable medium.

"Substantially identical" means a sequence having sufficient homology to hybridize under high stringency  
25 conditions and/or at least 90% identity at the nucleotide or amino acid sequence level to a sequence disclosed herein.

"Substantially purified" when used in reference to a protein or peptide means that the molecule has been largely, but not necessarily wholly, separated and purified  
30 from other cellular and non-cellular components. Typically a protein is substantially pure when it is at least about 60% by weight, free from other naturally occurring organic molecules. Preferably the purity is at least about 75%, more

-7-

preferably at least about 90%, and most preferably at least about 99% by weight pure.

"Selective hybridization" refers to hybridization under conditions of high stringency. Hybridization of nucleic acid molecules depends upon factors such as the degree of complementarity, stringency of hybridization conditions, and the length of hybridizing strands.

The term "stringency" relates to nucleic acid hybridization conditions. High stringency conditions disfavor non-homologous base pairing. Low stringency conditions have the opposite effect. Stringency may be altered, for example, by changes in temperature and salt concentration. Typical high stringency conditions comprise hybridizing at 50°C to 65°C in 5X SSPE and 50% formamide, and washing at 50°C to 65°C in 0.5X SSPE; typical low stringency conditions comprise hybridizing at 35°C to 37°C in 5X SSPE and 40% to 45% formamide and washing at 42°C in 1X-2X SSPE.

"SSPE" denotes a hybridization and wash solution comprising sodium chloride, sodium phosphate, and EDTA, at pH 7.4. A 20X solution of SSPE is made by dissolving 174 g of NaCl, 27.6 g of NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O, and 7.4 g of EDTA in 800 ml of H<sub>2</sub>O. The pH is adjusted with NaOH and the volume brought to 1 liter.

"SSC" denotes a hybridization and wash solution comprising sodium chloride and sodium citrate at pH 7. A 20X solution of SSC is made by dissolving 175 g of NaCl and 88 g of sodium citrate in 800 ml of H<sub>2</sub>O. The volume is brought to 1 liter after adjusting the pH with 10N NaOH.

"Virulence gene" as used herein means a gene from a pathogenic organism such as *S. pneumoniae* that is required for infection and/or pathogenicity in vivo. Some virulence

-8-

genes are induced during infection of a host; others are expressed exclusively during *in vivo* infection.

5 The *Streptococcus pneumoniae* genome contains about 2.2 million nucleotide base pairs and comprises about 2000 to 3000 ORFs and other genes. This invention provides, among other things, contiguous fragments, genes, and proteins from the *S. pneumoniae* genome (SEQ ID NO:1 through SEQ ID NO:228).

10 Strain differences in *S. pneumoniae* may be associated with nucleotide sequence differences in one or more of the genomic fragments disclosed herein. Sequences that are substantially identical to the sequences disclosed herein are intended to be within the scope of the invention.

15 The sequence fragments disclosed herein provide a wide variety of utilities. For example, the fragments may be used to identify regions of the *S. pneumoniae* genome that are expressed as proteins (*viz.* transcribed into mRNA). The genomic fragments disclosed herein can also be used to  
20 examine differential expression of *S. pneumoniae* genes under diverse environmental conditions, as occurs, for example, with the expression of virulence genes during *in vivo* infection of a host organism. Also contemplated by the invention are: (1) preparation of molecular hybridization  
25 probes for use in physical mapping, sequencing, mutagenesis, mutation analysis, (2) homology comparisons of the sequences disclosed herein with the genomes and ORFs of other organisms, (3) creation of specifically mutated strains of *S. pneumoniae* wherein the mutation is targeted to any site  
30 in the DNA sequence disclosed herein, (4) identification of *S. pneumoniae* promoters and other gene regulatory sequences, (5) identification of proteins and RNAs encoded by *S. pneumoniae*, (6) amplification of *S. pneumoniae* genes using the PCR, and (7) production of kits for isolating and

-9-

analyzing genes that are mutated in antibiotic resistant clinical isolates of *S. pneumoniae*.

#### Genome Analysis

5        In one embodiment, the invention comprises the ORFs and fragments thereof encoded by the nucleotide sequences disclosed herein. Some of the nucleotide sequences disclosed herein encode ORFs and fragments of ORFs (Table 1). The ORFs or fragments thereof were identified by translation of the  
10        nucleic acid sequences disclosed herein. The biological function of a protein disclosed in Table 1 was determined by homology comparison with known proteins from other organisms. A number of computer programs are available to assist in homology comparisons, for example Genemark  
15        (Borodovsky and McIninch, *Computers Chem.* 17(2), 123, 1993).

#### Computer-Related Applications

      The nucleotide and/or amino acid sequence information of this invention may be provided in a variety of media to  
20        facilitate use. In one embodiment the present invention comprises one or more of the sequences disclosed herein recorded on a computer readable medium. A variety of media are contemplated, for example, magnetic storage media such as floppy discs, hard disc storage, magnetic tape, and CD-  
25        ROM. A skilled artisan can readily adopt any presently known method for recording information on a computer readable medium to generate manufactures comprising the nucleotide or amino acid sequence information of the present invention. These embodiments are contemplated within the scope of this  
30        invention.

      The choice of a data storage structure will generally be based on the means chosen to access the stored information. A variety of data processor programs and formats can be used to store the sequence information of the

-10-

invention on computer readable medium. For example, the sequence can be represented in a word processing text file that is formatted in commercially available software such as WordPerfect and MicroSoft Word, or it can be represented in the form of a text only file such as ASCII.

Having *S. pneumoniae* genomic sequence information in a computer readable format enables a skilled artisan to access the information for a variety of purposes. For example, computer-assisted searching algorithms may be used to identify open reading frames, and ascertain biological function based on homology to known proteins from other organisms. Suitable algorithms for sequence comparisons include BLAST (Altschul et al., *J. Mol. Biol.* 215, 403-410, 1990) and BLAZE (Brutlag et al., *Comp. Chem.* 17, 203-207 (1993)). For identification of ORFs a number of commercially available software programs are suitable, such as FRAMES (Genetic Center Group, Madison, WI).

The genomic information of this invention in computer-readable form can be manipulated further using bioinformatics to identify the biological function of proteins encoded by ORFs as well as the cellular location of said proteins. The skilled artisan will recognize several computer-assisted algorithms for this purpose, for example, PSORT which is useful for determining the likely location of a protein within a cell (See K. Nakai & M. Kanehisa. "Expert system for predicting protein localization sites in Gram-negative bacteria", *Proteins: Structure, Function, and Genetics*, 11, 95-110 (1991)).

30

#### Open Reading Frames and Proteins

The invention also provides proteins encoded by the *S. pneumoniae* genome in substantially purified form (See Table 1). The proteins are classified herein as (1) Hypothetical,



-11-

(2) Cell wall biosynthetic, (3) External target, or (4) Minimal gene set proteins.

Cells that carry knockout mutations in proteins of the hypothetical class are nonviable. Loss of viability suggests that these proteins may be essential for viability. Two such proteins, whose genes map to contigs m014 and m016, correspond respectively to *Haemophilus influenzae* ORFs HI1146 and HI1648. Two other hypothetical proteins, yyaF and ywbL, correspond to a GTP binding protein and transcriptional regulator, respectively.

The proteins of this invention can be used to raise antibodies. Antibodies against the hypothetical class of proteins are especially attractive. In targeting presumptively essential cellular functions, antibodies against "hypothetical proteins" could have therapeutic or prophylactic applications. Additionally, the "hypothetical" proteins can be used to screen for agents that bind or otherwise interact with said proteins. Such agents could lead to the identification of new antibacterial agents.

Proteins classified in Table 1 as cell wall biosynthetic proteins, and external target proteins, were identified by homology with known proteins. These proteins are useful for identifying agents that bind and inhibit bacterial growth. Therefore, in another embodiment of the invention, the proteins of these classifications are prepared, preferably by recombinant means as described herein, substantially purified, and used in a screen to identify compounds that bind and/or inhibit the activity of said proteins. A variety of suitable screens are contemplated for this purpose. For example, the protein(s) can be labeled by known techniques such as radiolabeling or fluorescent tagging, or by labeling with biotin/avidin; thereafter binding of a test compound to a labeled protein

-12-

can be determined by any suitable means, well known to the skilled artisan.

The proteins categorized as "minimal gene set" are homologous to a set of highly conserved proteins found in other bacteria. The minimal gene set proteins are thought to be essential for viability, and are useful targets for the development of new antibacterial compounds.

#### DNA Chips and Applications

10       The nucleic acids disclosed herein, or subfragments thereof, may be arrayed on any suitable solid surface, thereby constructing a "chip." DNA chip hybridizations provide greater sensitivity than do conventional hybridization means, such as Southern hybridization or  
15 Northern hybridization. DNA chips are useful for a variety of purposes, for example, in mutation and gene expression analysis, and in probing the structure, function, and expression of the genome. This aspect of the invention relates to any one or more of the DNA fragments disclosed  
20 herein, wherein said fragments are attached to a solid support (i.e. "chip" or "DNA chip" or "Bio chip"). Attachment of a nucleic acid to a support can be, but is not necessarily, accomplished by chemical or enzymatic means.

In one embodiment, DNA fragments of this invention are  
25 arrayed onto a solid support as a means for assessing gene expression in *S. pneumoniae*. The DNA fragments attached to a chip may be of any size that is suitable for hybridization to other nucleic acid molecules such as cDNAs, genomic DNAs, or RNAs. Suitably-sized DNA fragments are from 10 nucleotide  
30 residues to approximately several thousand residues. The preferred length is about 50 to 500 nucleotides.

Analysis of gene expression using the chips of this invention is assessed by hybridization of a chip to RNA samples, or cDNA samples prepared from *S. pneumoniae* grown

-13-

under any suitable conditions. Preferred samples for hybridization to a chip comprise cDNA. Methods for preparing RNA or cDNA are well known in the art.

A variety of suitable methods are known for fixing DNA fragments to solid support matrices [See e.g. D. Stimpson et al. "Real-time detection of DNA hybridization and melting on oligonucleotide arrays by using optical wave guides" *Proc. Nat. Acad. Sci.* 92, 6379 (1995)] Preferred surfaces for producing a chip are glass or polystyrene. Convenient surfaces are microscope slides, or cover slips (Corning), treated with silicon or silane to minimize non-specific binding by DNA or proteins. Also suitable for this purpose are 96-well microtiter plates.

A light-directed method may be used for attaching oligonucleotides, enabling nucleotide synthesis directly on the solid surface using photolabile 5'-protected N-acyl-deoxynucleotide phosphoramidites and surface linker chemistry (See Pease et al. "Light-generated oligonucleotide arrays for rapid DNA sequence analysis" *Proc. Nat. Acad. Sci.* 91, 5022-5026, 1994). Alternatively, DNA fragments can be bound to a surface via interaction with a specific DNA binding protein. Any suitable DNA binding protein may be used, for example bacteriophage DNA binding proteins, Adenovirus binding protein, the *E. coli* lac-repressor protein, or  $\lambda$ -repressor protein. DNA binding proteins are attached to the surface of a chip by covalent chemical binding, essentially as described in U.S. Patent 5,561,071, the entire contents of which is incorporated by reference. The latter method requires that DNA fragments contain a recognition sequence that enables binding by the DNA binding protein. Specific sequences for a number of DNA binding proteins are known. Methods for incorporating specific binding sequences into the genomic DNA fragments disclosed herein are well known in the cloning arts.

-14-

DNA chip technology enables monitoring *S. pneumoniae* gene expression on a genome-wide level. This feature of the invention is particularly attractive for identifying (1) genes that are expressed or not expressed during the life cycle or infection cycle of *S. pneumoniae*, and (2) changes in gene expression that correlate with environmental change.

For example, virulence genes in *S. pneumoniae* can be identified by the DNA chip method disclosed herein. Identification of virulence genes in *S. pneumoniae* will provide new targets for developing novel antibiotics. For this aspect of the invention any suitable encapsulated strain of *S. pneumoniae* is introduced into a mouse, for example, by intraperitoneal injection, or by introduction directly into the lungs, or by any other suitable method. Approximately 2 days after infection a peripheral blood titre level is reached of about  $10^8$  *S. pneumoniae* cells/ml. Cells recovered from peripheral blood, or other suitable tissue, are used in identifying virulence genes. For this purpose, cDNAs are prepared from cells recovered from an *in vivo* infection and from cells grown *in vitro*. After labeling, the cDNAs are hybridized against the DNA chip(s) disclosed herein. Genomic fragments that hybridize to the *in vivo* probe but not to the *in vitro* probe identify candidate virulence genes.

Also contemplated by this aspect of the invention is a method for analyzing gene expression in *S. pneumoniae* cells grown or harvested from any desirable *in vitro* or *in vivo* environment, wherein said environment may include compounds whose effects on gene expression are to be determined.

In another embodiment, the present invention relates to a DNA bio-chip, useful for correlating DNA sequence with biological function. The bio-chip comprises an array of the genomic DNA fragments disclosed herein, or portions thereof, attached to the surface of any suitable solid support

-15-

material. The bio chip further comprises a layer of competent *S. pneumoniae* cells suspended over the DNA array in any suitable semi-solid medium such as agar or agarose. The cells suspended on the bio chip comprise known or  
5 unknown mutant strains, or they may be wild-type cells. The cell layer is in contact with the DNA matrix such that DNA on the chip can be taken up by the cells.

The bio-chip is useful for several purposes. For example, the bio-chip can be used to localize an unknown  
10 mutation to a specific region of the genome by complementation. The bio-chip enables correlating a phenotype with a genetic locus. For example, mutant cells harboring one or more mutations and having at least one screenable or selectable phenotype can be applied to a bio  
15 chip and subjected to an environment that allows for selection, or for screening by complementation. If said phenotype is the result of a chromosomal mutation or mutations that map to a genomic fragment present on the chip, DNA uptake by the cells and repair of the mutation by  
20 recombination will be identifiable by a suitable screen or selection.

In a preferred embodiment, the bio-chip is overlaid with competent *S. pneumoniae* cells. Methods for preparing competent cells are known (See e.g. LeBlanc et.al. *Plasmid*  
25 28, 130-145, 1992; Pozzi et al. *J. Bacteriol.* 178, 6087-6090, 1996).

Other embodiments of this aspect of the invention are contemplated. For example the genomic fragments disclosed herein could be prepared and dispensed into individual wells  
30 of a 96-well micro titre plate. Competent *S. pneumoniae* cells could then be added to the wells under conditions suitable for DNA uptake followed by plating onto any suitable selection or screening medium, for example an agar

-16-

plate containing suitable growth and/or selection/screening components.

#### Diagnostic Kits and Assays

5       The present invention further relates to kits and assays that can be used for rapid and efficient detection of *S. pneumoniae* cells. Also contemplated are kits for detecting mutations carried by *S. pneumoniae* cells. Kits of this nature are particularly attractive in the clinical  
10   environment where knowledge about the identity of a pathogen and/or of the basis for resistance to antibiotic treatments is essential for effective medical treatment. In the long term, knowledge of the mutations that lead to resistance will enable the design of new antibacterial agents.

15       A kit for detecting *S. pneumoniae* cells can be based on antibody recognition of *S. pneumoniae* specific antigens or epitopes, or by nucleic acid hybridization techniques for the detection of *S. pneumoniae* specific nucleic acid molecules.

20       A variety of embodiments are contemplated in this aspect of the invention. In one embodiment a kit is provided for detecting mutations in drug-resistant *S. pneumoniae*. For this purpose, DNA is prepared from a resistant isolate and from a wild-type strain. In a preferred embodiment, the  
25   polymerase chain reaction (i.e. PCR) is used to amplify DNA samples representing any one or all of the genomic fragments disclosed herein. The amplified DNAs from the mutant and wild-type cells are hybridized to a DNA chip having fixed thereon any one or more of the genomic fragments disclosed  
30   herein. Amplified DNA samples from the mutant and wild-type strain are labeled by any suitable means, for example using radioisotopes or fluorescent labeling. Hybridization of the amplified DNAs to the chip under conditions that can discriminate single or multiple base pair mismatches enables

-17-

the detection of differences between the mutant and wild-type samples. This method identifies a specific fragment of the genome that is altered in the mutant strain. The specific mutation can be determined by conventional DNA sequence analysis.

This aspect of the invention also relates to the detection of *S. pneumoniae* proteins in a sample using antibody molecules raised against any suitable ORF disclosed herein. Antibody detection methods are well known to those skilled in the art including, for example, a variety of radioimmunological assays. (See e.g. P. Tijssen, Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands, 1985).

Test samples suitable for use in this aspect of the invention include but are not limited to biological fluids such as sputum, blood, serum, plasma, urine, and to biopsy samples.

Skilled artisans will recognize that the disclosed method and reagents can be readily incorporated into a kit. For example, a kit would contain one or more receptacles comprising one or more of the following: PCR reagents, DNA chip reagents, labeling reagents, assorted buffers, and/or antibodies.

#### Production of Antibodies

The proteins of this invention and fragments thereof may be used in the production of antibodies. The term "antibody" as used herein describes antibodies, fragments of antibodies (such as, but not limited, to Fab, Fab', Fab<sub>2</sub>', and Fv fragments), and chimeric, humanized, veneered, resurfaced, or CDR-grafted antibodies capable of binding antigens of a similar nature as the parent antibody

-18-

molecule from which they are derived. The instant invention also encompasses single chain polypeptide binding molecules.

The production of antibodies, both monoclonal and polyclonal, in animals is well known in the art. See, e.g.,  
5 C. Milstein, Handbook of Experimental Immunology, (Blackwell Scientific Pub., 1986); J. Goding, Monoclonal Antibodies: Principles and Practice, (Academic Press, 1983). For the production of monoclonal antibodies the process begins with  
10 injecting a mouse, or other suitable animal, with an immunogen. The mouse is subsequently sacrificed and cells taken from its spleen are fused with myeloma cells, resulting in a hybridoma that can be cultured *in vitro*. Hybridomas are screened for clones that secrete a single antibody species, specific for the immunogen.

15 Chimeric antibodies, described in U.S. Patent No. 4,816,567, herein incorporated by reference, teaches methods and vectors for preparing chimeric antibodies. An alternative approach is provided in U.S. Patent No. 4,816,397, the entire contents of which is herein  
20 incorporated by reference. This patent teaches co-expression of heavy and light chains in the same host cell.

The method taught in U.S. Patent 4,816,397 has been further refined in European Patent Publication No. 0 239 400. The teachings of this publication are preferred for  
25 engineering monoclonal antibodies. In this technology the complementarity determining regions (CDRs) of a human antibody are replaced with the CDRs of a murine monoclonal antibody, thereby converting the specificity of the human antibody to the specificity of the murine antibody.

30 Single chain antibodies and libraries thereof provide yet another means for genetically engineering antibody molecules. (See, e.g. R.E. Bird, et al., *Science* 242:423-426 (1988); PCT Publication Nos. WO 88/01649, WO 90/14430, and WO 91/10737. Single chain antibody technology



-19-

involves covalently joining the binding regions of heavy and light chains thereby generating a single polypeptide chain having the binding specificity of an intact antibody molecule.

5       The antibodies contemplated by the present invention are useful in diagnostics, therapeutics, or in diagnostic/therapeutic combinations.

10       The proteins of this invention, or suitable fragments thereof, can be used to generate polyclonal or monoclonal antibodies, and various inter-species hybrids, or humanized antibodies, or antibody fragments, or single-chain antibodies. The techniques for producing antibodies are well known to skilled artisans. (See e.g. A.M. Campbell, Monoclonal Antibody Technology: Laboratory Techniques in  
15 Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam (1984); Kohler and Milstein, *Nature* 256, 495-497 (1975); Monoclonal Antibodies: Principles & Applications Ed. J.R. Birch & E.S. Lennox, Wiley-Liss, 1995.

20       A protein or peptide to be used as an immunogen may be administered in an adjuvant by subcutaneous or intraperitoneal injection into, for example, a mouse or a rabbit. For the production of monoclonal antibodies, spleen cells from immunized animals are removed, fused with myeloma cells, such as SP2/0-Ag14 cells, and allowed to become  
25 monoclonal antibody producing hybridoma cells in the manner known to the skilled artisan. Hybridomas that secrete the desired antibody molecule can be screened by a variety of well known methods, for example ELISA assay, western blot analysis, or radioimmunoassay (Lutz et al. *Exp. Cell Res.*  
30 175, 109-124 (1988); Monoclonal Antibodies: Principles & Applications Ed. J.R. Birch & E.S. Lennox, Wiley-Liss, 1995).

For some applications it is desirable to have an antibody labeled in some fashion. Procedures for labeling antibody molecules with radioisotopes, affinity labels, such

-20-

as biotin or avidin, enzymatic labels, for example horseradish peroxidase, and fluorescent labels such as FITC or rhodamine, are widely known (See e.g. Enzyme-Mediated Immunoassay, Ed. T. Ngo, H. Lenhoff, Plenum Press 1985; 5 Principles of Immunology and Immunodiagnostics, R.M. Aloisi, Lea & Febiger, 1988).

Labeled antibodies are useful for a variety of diagnostic applications. In one embodiment, the present invention relates to the use of labeled antibodies to detect 10 the presence of *S. pneumoniae* cells and proteins. Also contemplated are applications that use antibodies, preferably single chain antibodies, directed against a *S. pneumoniae* protein. Proteins identified as "external targets" are preferred for the generation of single chain 15 antibodies. Single chain antibody libraries directed against *S. pneumoniae* surface proteins and cell wall proteins can be produced by applying the phage display technique to crude membrane preparations. Antibodies that recognize and bind to external target proteins and/or cell wall proteins could be 20 used as therapeutic agents to inhibit the growth of *S. pneumoniae*. Alternatively, the antibodies could be used in a screen to identify potential inhibitors of an external target protein. For example, in a competitive displacement assay, an antibody or compound to be tested is labeled by 25 any suitable method. Competitive displacement of an antibody from an antibody-antigen complex by a test compound provides a means to identify new antibacterial compounds.

#### Protein Production Methods

30 The present invention relates further to substantially purified proteins encoded by the ORFs disclosed herein (SEQ ID NO:87 through SEQ ID NO:228).

Skilled artisans will recognize that proteins can be synthesized by different methods, for example, chemical

-21-

methods or recombinant methods, as described in U.S. Patent 4,617,149, hereby incorporated by reference.

The principles of solid phase chemical synthesis of polypeptides are well known in the art and may be found in general texts relating to this area. See, e.g., H. Dugas and C. Penney, Bioorganic Chemistry (1981) Springer-Verlag, New York, 54-92. Peptides may be synthesized by solid-phase methodology utilizing an Applied Biosystems 430A peptide synthesizer (Applied Biosystems, Foster City, CA) and synthesis cycles supplied by Applied Biosystems. Protected amino acids, such as t-butoxycarbonyl-protected amino acids, and other reagents are commercially available from many chemical supply houses.

The proteins and peptides of the present invention can also be made by recombinant DNA methods. Recombinant methods are preferred if a high yield is desired. Recombinant methods involve expressing a cloned ORF/gene in a suitable host cell. A gene is introduced into a host cell by any suitable means, well known to those skilled in the art. While chromosomal integration of a cloned gene is within the scope of the present invention, it is preferred that a cloned gene be maintained extra-chromosomally, as part of a vector wherein the gene is in operable-linkage to a constitutive or inducible promoter.

Recombinant methods are also useful in overproducing a membrane-bound or membrane-associated protein. In some cases, membranes prepared from recombinant cells that overexpress such proteins provide an enriched source of the protein. Such membranes are useful for evaluating the function of the protein and/or for evaluating inhibitors of the protein.

-22-

Expressing Recombinant Proteins in Procaryotic and  
Eucaryotic Host Cells

Procaryotes are generally used for cloning DNA sequences and for constructing vectors. For example, the  
5 *Escherichia coli* K12 strain 294 (ATCC No. 31446) is particularly useful for expression of foreign proteins. Other strains of *E. coli*, bacilli such as *Bacillus subtilis*, enterobacteriaceae such as *Salmonella typhimurium* or *Serratia marcescans*, various *Pseudomonas* species may also be  
10 employed as host cells in cloning and expressing the recombinant proteins of this invention. Also contemplated are various strains of *Streptococcus* and *Streptomyces*.

For effective expression of a recombinant protein a gene or ORF may be linked to a known promoter sequence.  
15 Suitable bacterial promoters include *b* -lactamase [e.g. vector pGX2907, ATCC 39344, contains a replicon and *b* -lactamase gene], lactose systems [Chang et al., *Nature* (London), 275:615 (1978); Goeddel et al., *Nature* (London), 281:544 (1979)], alkaline phosphatase, and the tryptophan  
20 (*trp*) promoter system [vector pATH1 (ATCC 37695)] designed for the expression of a *trpE* fusion protein. Hybrid promoters such as the *tac* promoter (isolatable from plasmid pDR540, ATCC-37282) are also suitable. Promoters for use in bacterial systems also will contain a Shine-Dalgarno  
25 sequence operably linked to the DNA encoding the desired polypeptides. These examples are illustrative rather than limiting.

A variety of mammalian cell systems and yeasts are also suitable host cells. The yeast *Saccharomyces*  
30 *cerevisiae* is a commonly used eucaryotic microorganism. Other yeasts such as *Kluyveromyces lactis* are also suitable. For expression of recombinant genes in *Saccharomyces*, the plasmid YRp7 (ATCC-40053), for example, may be used. See, e.g., L. Stinchcomb, et al., *Nature*, 282:39 (1979); J.

-23-

Kingsman et al., *Gene*, 7:141 (1979); S. Tschemper et al., *Gene*, 10:157 (1980). Plasmid YRp7 contains the TRP1 gene that provides a selectable marker in a *trp1* mutant.

5 Purification of Recombinantly-Produced Protein

An expression vector carrying an ORF of the present invention is transformed or transfected into a suitable host cell using standard methods. Cells which contain the vector are propagated under conditions suitable for expression of the encoded protein. If the gene is under the control of an inducible promoter then suitable growth conditions would incorporate the appropriate inducer. The recombinantly-produced protein may be purified from cellular extracts of transformed cells by any suitable means.

15 In a preferred process for protein purification a gene/ORF is modified at the 5' end, or some other position, to incorporate a plurality of histidine residues at the amino terminus of the encoded protein. The "histidine tag" produced thereby enables a single-step protein purification method referred to as "immobilized metal ion affinity chromatography" (IMAC), essentially as described in U.S. Patent 4,569,794, hereby incorporated by reference. The IMAC method enables rapid isolation of substantially pure protein starting from a crude cellular extract.

25 As skilled artisans will recognize, the proteins of the invention can be encoded by a multitude of different nucleic acid sequences owing to the degeneracy of the genetic code. The present invention further comprises these alternate nucleic acid sequences.

30 The ribonucleic acid compounds of the present invention may be prepared using the polynucleotide synthetic methods discussed *supra*, or they may be prepared enzymatically using RNA polymerase to transcribe a DNA template.

-24-

The most preferred systems for preparing the ribonucleic acids of the present invention employ the RNA polymerase from the bacteriophage T7 or the bacteriophage SP6. These RNA polymerases are highly specific, requiring  
5 the insertion of bacteriophage-specific sequences at the 5' end of the template to be transcribed. See, J. Sambrook, et al., *supra*, at 18.82-18.84.

This invention also provides nucleic acids, RNA or DNA, which are complementary to the sequences disclosed  
10 herein.

The present invention also provides probes and primers useful for a variety of molecular biology techniques including, for example, hybridization screens of genomic or subgenomic libraries, detection and quantification of mRNA  
15 species as a means to analyzing gene expression, and amplification of any region of the *Streptococcus pneumoniae* genome disclosed by the sequences herein. A nucleic acid compound is provided comprising any of the sequences disclosed herein, or a complementary sequence thereof, or a  
20 fragment thereof, which is at least 15 base pairs in length, and which will hybridize selectively to *Streptococcus pneumoniae* DNA or mRNA. Preferably, the 15 or more base pair compound is DNA. A probe or primer length of at least 15 base pairs is dictated by theoretical and practical  
25 considerations. See e.g. B. Wallace and G. Miyada, "Oligonucleotide Probes for the Screening of Recombinant DNA Libraries," In Methods in Enzymology, Vol. 152, 432-442, Academic Press (1987).

The probes and primers of this invention can be  
30 prepared by methods well known to those skilled in the art (See e.g. Sambrook et al. *supra*). In a most preferred embodiment these probes and primers are synthesized by the polymerase chain reaction (PCR).

-25-

The present invention also relates to recombinant DNA cloning vectors and expression vectors comprising the nucleic acids of the present invention. Preferred nucleic acid vectors are those which comprise DNA. The skilled  
5 artisan understands that choosing the most appropriate cloning vector or expression vector depends on a number of factors including the availability of restriction enzyme sites, the type of host cell into which the vector is to be transfected or transformed, the purpose of the transfection  
10 or transformation (e.g., stable transformation as an extrachromosomal element, or integration into the host chromosome), the presence or absence of readily assayable or selectable markers (e.g., antibiotic resistance and metabolic markers of one type and another), and the number  
15 of gene copies desired in the host cell.

Vectors suitable to carry the nucleic acids of the present invention comprise RNA viruses, DNA viruses, lytic bacteriophages, lysogenic bacteriophages, stable  
20 bacteriophages, plasmids, viroids, and the like. The most preferred vectors are plasmids.

Host cells harboring the nucleic acids disclosed herein are also provided by the present invention. A preferred host is *E. coli* which has been transfected or transformed with a vector that comprises a nucleic acid of  
25 the present invention.

The present invention also provides a method for constructing a recombinant host cell capable of expressing an ORF disclosed herein, said method comprising transforming or otherwise introducing into a host cell a recombinant DNA  
30 vector that comprises an isolated DNA sequence which encodes said ORF. The preferred host cell is any strain of *E. coli* which can accommodate high level expression of an exogenously introduced gene. Transformed host cells are cultured under conditions well known to skilled artisans such that said ORF

-26-

is expressed, thereby producing the encoded protein in the recombinant host cell.

For the purpose of discovering new inhibitors of cell wall biosynthesis, it would be desirable to determine agents that inhibit enzymes required for synthesis of the cell wall and/or agents that interact with membrane proteins. A method for identifying compounds that interact with such enzymes and membrane proteins comprises contacting said proteins with a test compound and monitoring an interaction and/or inhibition by any suitable means.

The instant invention provides a screening system for compounds that interact with membrane proteins of this invention, said screening system comprising the steps of:

- a) preparing a membrane protein, or membranes enriched in said protein;
- b) exposing the protein source of (a) to a test compound; and
- c) quantifying the interaction of said protein with said compound by any suitable means.

20

The screening method of this invention may be adapted to automated procedures such as a PANDEX® (Baxter-Dade Diagnostics) system, allowing for efficient high-volume screening of compounds.

25

In a typical screening protocol, a protein to be tested is prepared as described herein, preferably using recombinant DNA technology. A test compound is introduced into a reaction vessel containing said protein. The reaction/interaction of said protein and said compound is monitored by any suitable means. For example, a radioactively-labeled or chemically-labeled compound or

30



-27-

protein may be used. Specific association between a test compound and protein is monitored by any suitable means.

The following examples more fully describe the present invention. Those skilled in the art will recognize  
5 that the particular reagents, equipment, and procedures described are merely illustrative and are not intended to limit the present invention in any manner.

## EXAMPLE 1

10 Vector for Expressing *S. pneumoniae* ORF in a Host Cell

An expression vector suitable for expressing a *S. pneumoniae* gene or fragment thereof in a variety of procaryotic host cells, such as *E. coli*, is easily made. A suitable parent vector contains an origin of replication  
15 (Ori), a marker for selecting transformants, for example, an ampicillin resistance gene (Amp), and further comprises suitable transcriptional and translational signals, for example, the T7 promoter and T7 terminator sequences, in operable-linkage to a *S. pneumoniae* coding region. For  
20 example, pET11A (obtained from Novogen, Madison WI) is linearized by restriction with endonucleases NdeI and BamHI. Linearized pET11A is ligated to a DNA fragment bearing NdeI and BamHI sticky ends and comprising a coding region for a *S. pneumoniae* ORF.

25 The ORF used in this construction may be modified at the 5' end (amino terminus of encoded protein or peptide) to simplify purification of the encoded protein or peptide. For this purpose, an oligonucleotide encoding 8 histidine residues is inserted after the transcriptional and  
30 translational start sites. Placement of the histidine residues at the amino terminus of the encoded protein enables the IMAC one-step protein purification procedure.

Example2

-28-

Recombinant Expression and Purification of a Protein Encoded  
by a *S. pneumoniae* ORF

An expression vector that carries an ORF from the *S. pneumoniae* genome, as disclosed in Example 1, and which  
5 ORF is operably-linked to an expression promoter, is transformed into *E. coli* BL21 (DE3) (*hsdS gal lcIts857 ind1Sam7nin5lacUV5-T7gene 1*) using standard methods. Transformants, selected for resistance to ampicillin, are chosen at random and tested for the presence of the vector  
10 by agarose gel electrophoresis using quick plasmid preparations. Colonies that contain the vector are grown in L broth and the protein produced by the vector-borne ORF is purified by IMAC, essentially as described in US Patent 4,569,794.

15 Briefly, the IMAC column is prepared as follows. A metal-free chelating resin (e.g. Sepharose 6B IDA, Pharmacia) is washed in distilled water to remove preservatives and then infused with a suitable metal ion [e.g. Ni(II), Co(II), or Cu(II)] by adding a 50mM metal  
20 chloride or metal sulfate aqueous solution until about 75% of the interstitial spaces of the resin are saturated with colored metal ion. The column is then ready to receive a crude cellular extract containing the recombinant protein product.

25 Unbound proteins and other materials are removed by washing the column with any suitable buffer, pH 7.5. Bound protein is eluted in any suitable buffer at pH 4.3, or preferably with an imidazole-containing buffer at pH 7.5.

30

Example 3

DNA Chip Production

Any one or more of the *S. pneumoniae* genome DNA fragments disclosed herein, or fragments thereof, are arrayed onto a solid support. It is preferred that fragments be in

-29-

the size range of 14 base pairs to 500 base pairs. The DNA samples are most conveniently synthesized by PCR using standard methods to amplify regions disclosed by the genomic sequences herein. The method of Schena et al. is used to spot about 1 ng to 10 ng of a DNA sample onto glass microscope slides that have been treated with poly-L-lysine (M. Schena et al. "Quantitative monitoring of gene expression patterns with a complementary DNA microarray" Science, 270, 467-470, 1995). After spotting DNA samples onto the chip and air-drying, the chips are rehydrated by incubation for about 2 hours in a humid chamber. Chips are then placed at 100° C for 1 minute, rinsed in 0.1% SDS, and treated with 0.05% succinic anhydride in 50% 1-methyl-2-pyrrolidinone and 50% boric acid.

#### Example 4

##### S. pneumoniae Gene Expression Analysis using DNA Chips

RNA prepared from cells grown under any desirable conditions is used to prime cDNA synthesis by reverse transcription, using methods well known to the skilled artisan (See e.g. Molecular Cloning, 2d Ed. J. Sambrook, E. Fritsch, T. Maniatis, 1989). For example, total RNA of strain R6 is prepared according to the method of Logeman et.al., (Analytical Biochemistry, 1987, 163, 16-20) using guanidine hydrochloride. After ethanol precipitation, the total RNA is dissolved in a buffered solution such as Tris-EDTA (TE). Complementary DNA's are synthesized with the aid of the StrataScript RT-PCR kit (Stratagene, Inc.) in accordance with the supplier's recommendations (See Schena et al. *Id.*). Briefly, a 50 ul reaction contains about 0.1 ug/ul of RNA. First strand synthesis is primed using random primers, 1X first strand buffer, 0.03 U/ul ribonuclease block, 500 uM dATP, 500 uM dTTP, 40 uM dGTP, 40 uM fluorescein-12-dCTP (New England Nuclear), and 0.03 U/ul

-30-

reverse transcriptase. Reactions are incubated for 60 minutes at 37° C, precipitated with ethanol, and resuspended in 10 ul TE pH 8. Samples are heated for 3 minutes at 94° C and chilled on ice. The RNA is degraded by adding 0.25 ul of 10 N NaOH, followed by a 10 minute incubation at 37° C. The samples are neutralized with 2.5 ul of 1M Tris-HCl, pH 8 and 0.25 ul of 10 N HCl. After ethanol precipitation, the nucleic acid pellet is washed and dried *in vacuo*.

Prior to hybridization, DNA chips prepared as in Example 3 are denatured by heating to 90°C for 2 minutes. Hybridization reactions contain about 1 ul of fluorescently-labeled cDNA, and 1 ul of hybridization buffer (10x SSC and 0.2% SDS). Probe mixtures are transferred to the surface of the chip, covered with a cover slip, and incubated for 18 hours at 65° C. Chips are washed 5 minutes at room temperature in 1X SSC, 0.1% SDS, then for 10 minutes at room temperature in 0.1X SSC, 0.1% SDS. After hybridization, chips are scanned with a laser-scanning device.

-31-

## Example 5

A DNA Bio Chip for mutation analysis

Duplicate DNA chips are prepared as in Example 3. Each chip is overlaid with *S. pneumoniae* cells in a semi-solid medium, wherein said cells carry a temperature-sensitive (ts) mutation in a gene required for autolytic activity (Lyt<sup>-</sup>). This mutation leads to resistance to lysis at 37° C, but sensitivity to lytic treatments at 30° C.

*S. pneumoniae* strain cwl is resistant to lysis by detergent and penicillin when grown at 37° C, but remains sensitive when grown at 30° C (cwl is derived from strain R6; See P. Garcia et al. "Mutants of Streptococcus pneumoniae that contain a temperature-sensitive autolysin" *J. Gen. Microbiol.* 132, 1401-05, 1986). Strain cwl is grown at 30° C and competent cells are prepared according to any suitable method (e.g. LeBlanc et.al. *Plasmid* 28, 130-145, 1992; Pozzi et al. *J. Bacteriol.* 178, 6087-6090, 1996). Competent cwl cells are harvested by centrifugation and resuspended at about 10<sup>5</sup> cells per ml in 1% melted agar supplemented with 0.1% (w/v) yeast extract (Difco) and containing 1% to 2% Triton X-100. Approximately 100 ul to 500 ul of the cell mixture is deposited per square centimeter onto the bio chip by pipetting onto the chip surface. After solidification of the agar layer, one of the bio-chips is incubated at 37° C and the other at 30° C. Cells that take up a complementing genomic DNA fragment from the chip surface will be lysed at both 30° C and 37° C, while non-complemented cells are lysed only at 30° C. Cells that are complemented by the bio-chip are recognizable by this phenotypic difference and can be further purified by well known methods.

TABLE 1  
S. pneumoniae Genes and  
proteins

Contig ID	Associated ORF(s)	Nucleotide (Amino Acid SEQ ID NO)	Gene category	Comments
j001	yvaF	1 (123)	hypothetical	yvaF= GTP-binding protein
j002	obg	2 (124)	external	obg=GTP-binding protein
j003	lspA; ywbI	3 (126) 3 (125)	cell wall synth; hypothetical	lsp=prolipoprotein signal peptidase ywbI=hypothetical transcriptional regulator
j004	pstA	4 (127)	hypothetical external	pstA=high-affinity phosphate-specific transport
j005	pilB	5 (128)	external	pilB=fimbrial transcription regulation repressor
j008	pgsA	6, 111	external	pgsA=phosphatidylglycerophosphate synthase
j009	potC; potF	7 (130) 7 (131)	external	potC, potF=spermidine/putrescine transport system permease proteins
j010	pbuX; uraA(pyrP)	8, 9 (132) (133)	external ;	pbuX=xanthine permease uraA=uracil permease
j012	nanB/A	10 (134)	external	nanB/A=neuraminidase

-33-

j013	spoIIIEB	11 (135)	external	spoIIIEB=sporulation protein
j014	licC	12 (136)	external	licC=(OM) lipopolysaccharide phase variation protein
m001	unk	13 (137)	hypothet ical	
m002	unk	14 (138)	hypothet ical	
m003	unk	15 (139)	hypothet ical	
m004	unk	16, 95 (140)	hypothet ical	
m005	unk	17 (141)	hypothet ical	
m007	murZ	18, 103	hypothet ical	
m008	unk	19 (142)	hypothet ical	
m011	unk, unk unk unk	20 (143) 20 (144) 20 (145) 20 (146)	hypothet ical hypothet ical	
m012	abctran, aminoatrans, glutrans	21 (147), 22 (148), 23(149), 24 (150) 25 (151) 26, 121	external	transport proteins
m014	unk (HI1146)		hypothet ical	

-34-

m015	ftsH hgprt	27, 117 (153)	external	putative ATPase involved in membrane functions
m016	unk, unk (HI1648)	28, 109 (154, 155)	hypothet ical	
m017	nusG	29 (156)	min set w/o euk match	transcription elongation factor
m018	unk	30 (157)	hypothet ical	
m019	era	31, 91	external	membrane associated GTP binding protein
m021	yhgf	32 (158)	external	hypothetical
m022	rpoA, secY KAD	33 (158) 115 (159)	external min set w/o euk match	rpoA encodes alpha subunit of RNA polymerase, secY is a translocator of proteins (multispanning membrane protein); KAD = adenylate cyclase CTP synthetase
m023	pyrG	34 (161)	external	amidase
m024	amiF	35 (162)	external	
m027	unk	36 (163)	external	
m028	spot	37 (164)	external	guanosine 5'- diphosphate, 3'- diphosphate pyrophosphatase



-35-

m030	mray	38, 99	cell wall syn/memb rane int. cell wall syn/memb rane int. cell wall syn/memb rane int.	UDP-N <sub>2</sub> acetylmuramoyl- pentapeptide:undecapre nyl_P)4 phosphatase
m031	rodA	39, 113		
m034	murB	40 (165)		UDP-N <sub>2</sub> - acetylglucosaminyl-3- enolpyruvate reductase
m036	ftsY	41, 119	int. cell wall syn/memb rane int.	function in cell division
p001	nrfE; lacRrepressor nrdF	42 (166) (167) (168)	minimal set	nrfE: Nitrite reduction, formate dependent, nrdF: nucleotide reductase
p002	ACPSyn; fabB; ACPtrans; pepN; unk unk ligase	43, 44, 45 (172) (170) (171) (173) (169) (174) 46, 47, 97	external      minimal set minimal	fabB: fatty acid biosynthesis, pepN: aminopeptidase N  ligase
p004	trigger factor	48 (176)		trigger factor

-36-

p005	ptsI	49, 50 (177, 178)	set external	PRP-protein PTS enzyme I
p006	ftsZ	51, 105	external	filamentous temperature sensitivity, cell septation
p007	unk; unk; ftsE; unk	52 (180) (181) (179) (182)	cell wall synth.	ftsE: filamentous temperature sensitivity, ATP- binding protein
p008	glpF;	53 (183)	external	glpF: glycerol facilitator
p009	grpE;	54, 107	external	grpE: nucleotide exchange factor
p010	unk; unk unk unk unk unk unk	55 (186) (187) (188) (189) (190) (191) (192)	cell wall synth (murI region)	
p012	topA	56 (193)	minimal set	DNA topoisomerase I
p013	doeD	57 (194)	minimal set	purine-nucleoside phosphorylase
p014	rfbX; lic-1 operon	58, (195) 59, 60 (196-198)	cell wall synth.	rfbX: hydrophobic protein involved in O- antigen assembly

-37-

p016	ATPase;	61, 62, 63, 64 (199-202)	external	cation transporting ATPase
p017	livH; braC	65 (204) 65 (203)	external	livH: leucine, isoleucine and valine transport
p018	rimI; gcp	66 (205) 93	external	rimI: modification of 30S ribosomal subunit, gcp: essential gene of unknown function
Punk01	unk;	67 (206-207)	hypothet ical	
Punk02	unk	68 (208)	hypothet ical	
Punk03	unk	69 (209)	hypothet ical	
Punk04	unk	70 (210)	hypothet ical	
Punk05	unk;	71 (211-212)	hypothet ical	
Punk06	unk;	72 (213-214)	hypothet ical	
Punk08	unk	73 (215)	hypothet ical	
Punk09	unk;	74 (216-217)	hypothet ical	
Punk10	unk	75 (218)	hypothet ical	
Punk12	unk	76 (219)	hypothet ical	
s001	AtpF	77 (220)	minimal set	ATP synthase B chain
s002	DnaG	78, 89	minimal set	DNA primase

s003	EL-G; polC	79 (222) (221)	minimal set	elongation factor G; DNA polymerase III
s004	FAD	80 (223)	minimal set	FAD synthetase
s005	SmpB	81 (224)	minimal set	regulatory protein LPS-heptosyl-2- transferase, small prot.
s006	DnaA	82 (225)	minimal set	chromosomal replication initiator
s008	DnaB	83, 87	minimal set	replicative DNA helicase
s009	LgtP	84 (226)	minimal set	prolipoprotein diacylglycerol transferase
s010	UDPgalactose Epimerase glycosyl transferase	85 (227) 86 (228)	cell wall synth	UDPgalactose Epimerase
	dpj-acps	101		

-39-

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Baltz, Richard H.  
Burgett, Stanley G.  
DeHoff, Bradley S.  
10 Jaskunas Jr., Stanley R.  
Mills, Bradley J.  
Norris, Franklin H.  
Peery, Robert B.  
Rosteck Jr., Paul R.  
Skatrud, Paul L.  
15 Smith, Michele C.  
Rockey, Pamela K.  
Young-Bellido, Michele
- (ii) TITLE OF INVENTION: Streptococcus Pneumoniae DNA Sequences
- 20 (iii) NUMBER OF SEQUENCES: 122
- (iv) CORRESPONDENCE ADDRESS:  
25 (A) ADDRESSEE: Eli Lilly and Company  
(B) STREET: Lilly Corporate Center  
(C) CITY: Indianapolis  
(D) STATE: Indiana  
(E) COUNTRY: U.S.  
30 (F) ZIP: 46285
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
35 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER:  
40 (B) FILING DATE:  
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Webster, Thomas D.  
(B) REGISTRATION NUMBER: 39,872  
45 (C) REFERENCE/DOCKET NUMBER: X-11162
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 317-276-3334
- 50 (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1267 base pairs  
55 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 60 (iii) HYPOTHETICAL: NO

-40-

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

10	ATGGTGGGAAG TTCCAGATGA ACGCCTACAA AAATAACTG AAATGATAAC TCCTAAAAAG	60
	ACAGTTCCCA CAACATTTGA ATTTACAGAT ATTGCAGGGA TTGTAAAAGG AGCTTCAAAA	120
	GGAGAAGGGC TAGGGAATAA ATTCTTGGCC AATATTCGTG AAGTAGATGC GATTGTTCAC	180
15	GTAGTTCGTG CTTTTGATGA TGAAAATGTA ATGCGCGAGC AAGGACGTGA AGACGCCTTT	240
	GTAGATCCAC TTGCAGATAT TGATACAATT AATCTGGAAT TAATCTTGC TGACTTAGAA	300
	TCAGTGAACA AACGATATGC GCGTGTAGAA AAGATGGCAC GTACGCAAAA AGATAAAGAA	360
20	TCAGTAGCAG AATTCAATGT TCTTCAAAAG ATTAAACCAG TCCTAGAAGA CGGGAAATCA	420
	GCTCGTACCA TTGAATTTAC AGATGAGGAA CAAAAGGTTG TCAAAGGTCT TTTCCTTTTG	480
25	ACGACTAAAC CAGTTCTTTA TGTAGCTAAT GTGGACGAGG ATGTGGTTTC AGAACCTGAC	540
	TCTATCGACT ATGTCAAACA AATTCGTGAA TTTGCAGCGA CAGAAAATGC TGAAGTAGTC	600
	GTTATTTCTG CGCGTGCTGA GGAAGAAATT TCTGAATTGG ATGATGAAGA TAAAAAGAG	660
30	TTTCTTGAAG CCATTGGTTT GACAGAATCA GGTGTAGATA AGTTGACGCG TGCAGCTTAC	720
	CACTTGCTTG GATTGGGAAC TTA CTTCACA GCTGGTGAAA AAGAAGTTCG CGCTTGGACT	780
35	TTCAAACGTG GTATGAAGGC TCCTCAAGCA GCTGGTATTA TCCACTCAGA CTTTGAAAAA	840
	GGCTTTATTC GTGCAGTAAC CATGTCATAT GAAGATCTAG TGAAATACGG ATCTGAAAAG	900
	GCCGTAAAAG AAGCTGGACG CTGCGTGAA GAAGGAAAAG AATATATCGT TCAAGATGGC	960
40	GATATCATGG AATTCGCTT TAATGTCTAA AAATTAATAA ATGGTGTCAA TTAGGTTGGA	1020
	AAAAAATTCC AACCTTTTG GCTTTTGAAA GGA AAAATAA ATGACCAAAT TACTTGTAGG	1080
45	TTTGGGAAAT CCAGGGGATA AATATTTTGA AACAAACACA ATGTTGGTTT TATGTTGATT	1140
	GATCAACTAG CGAAGAAACA GAATGTCACT TTTACACACG ATAAGATATT TCAAGCTGAC	1200
	CTAGCATCCT TTTTCCTAAA TGGAGAAAAA ATTTATCTGG GTTAAACCAA CGACCTTTAT	1260
50	GGATTGA	1267

(2) INFORMATION FOR SEQ ID NO:2:

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1255 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

60

(ii) MOLECULE TYPE: DNA (genomic)

-41-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10

TGGTCCGTGG TGCTGAGGAC CCTTAGAGTT CGAGTACCAC AAGGTACCGA CTGTTCGTGA	60
TGCGGAGCTG GCAAGGTTTT AACAGATTTG ATTGAACATG GGCAAGAATT TATCGTTGCC	120
CACGGTGGTC GTGGTGGACG TGGAAATATT CGTTTCGCGA CACCAAAAAT CCTGCACCGG	180
AATCTCTGAA AATGGAGAAC CAGGTCAGGA ACGTGAGTTA CAATTGGAAC TAAAAATCTT	240
GGCAGATGTC GGTTTAGTAG GATTCCCATC TGTAGGGAAG TCAACACTTT TAAGTGTTAT	300
TACCTCAGCT AAGCCTAAAA TTGGTGCCTA CCACTTTACC ACTATTGTAC CAAATTTAGG	360
TATGGTTCGC ACCCAATCCA GGTGAATCCT TTGCAGTAGC CGACTTGCCA GGTTTGATTG	420
AAGGGGCTAG TCCAAGGTGT TGGTTTGGGA ACTCAGTTCC TCCGTCACAT CGAGCGTACA	480
CGTGTTATCC TTCACATCAT TGATATGTCA GCTAGCGAAG GCCGTGATCC ATATGAGGAT	540
TACCTAGCTA TCAATAAAGA GCTGGAGTCT TACAATCTTC GCCTCATGGA GCGTCCACAG	600
ATTATTGTAA CTAATAAGAT GGACATGCCT GAGAGTCAGG AAAATCTTGA AGAATTTAAG	660
AAAAAATTGG CTGAAAATTA TGATGAATTT GAAGAGTTAC CAGCTATCTT CCCAATTTCT	720
GGATTGACCA AGCAAGGTCT GGCAACACTT TTAGATGCTA CAGCTGAATT GTTAGACAAG	780
ACACCAGAAT TTTTGCTCTA CGACGAGTCC GATATGGAAG AAGAAGTTTA CTATGGATTT	840
GACGAAGAAG AAAAAGCCTT TGAAATTAGT CGTGATGACG ATGCGACATG GGTACTTTCT	900
GGTGAAAAAC TCATGAAACT CTTTAATATG ACCAACTTTG ATCGTGATGA ATCTGTCATG	960
AAATTTGCCC GTCAGCTTCG TGGTATGGGG GTTGATGAAG CCCTTCGTGC GCGTGGAGCT	1020
AAAGATGGGG ATTTGGTCCG CATTGGTAAA TTTGAGTTTG AATTTGTAGA CTAGGAGACT	1080
GGTATGGGAG ATAAACCGAT ATCTTTCCGA GATGCGGATG GTAATTTTGT TTCCGCCGCA	1140
GACGTTTGGA ATGAAAAGAA ATTGGAAGAA CTATTTAATC GTCTCAATCC AAATCGTGCC	1200
TTGAGATTGG CACGAACCTAC AAAGGAAAAT CCATCTCAGT AAAGAAGCTA AAAAA	1255

(2) INFORMATION FOR SEQ ID NO:3:

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1609 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

60

(ii) MOLECULE TYPE: DNA (genomic)

-42-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

10

TTACCCATCG CATGACTAAA AATCTCTACT ATCCAATACT AGTTCATATT CTCATCAATA	60
TCACTGCCTT CTGGGATGTT TGGTTACTCC TATTTTCAGG AAGTTAGCTT ACTAAAAAAA	120
15 TGTCGGAATT TTCCGGCATT TTCTTTTTTC ACAAATAGTC AACGTTTTTC TTTCCGATAC	180
TGAAGTGGTG TGTAGCCACT TATTTTTTTG AATTGATTTT GAAAATAAGA TTGGCGTGAG	240
20 AAAGGCAGAT AGTGAAGATA GTTAAGAAGA ATAGGATGTT CTTTTTCCT TTTTGAAAA	300
CTTCTAAAAT ATGGTATAAT GAAAAGATAA AGAAGTTGGG GGTAGAAGAT GAACATTCAA	360
CAATTACGCT ATGTTGTGGC TATTGCCAAT AGTGGTACTT TTCGTGAAGC TGCTGAAAAG	420
25 ATGTATGTTA GTCAGCCGAG TCTGTCTATT TCTGTTCTG ATTTGAAAA AGAGTTGGGC	480
TTTAAGATTT TCCGTCGGAC CAGCTCAGGG ACTTCTTGA CCCGTCGTGG GATGGAATTT	540
TATGAAAAAG CGCAAGAATT GGTAAAGGA TTTGATATTT TTCAAATCA GTATGCCAAT	600
30 CCTGAAGAAG AAAAAGATGA ATTTCCGTT GCTAGCCAGC ACTATGACTT CTTACCACCA	660
ACTATTACGG CCTTTTCAGA GCGCTATCCT GACTATAAGA ACTTCCGTAT TTTTGAATCA	720
35 ACTACTGTTC AAATATTAGA TGAAGTGGCG CAAGGGCATA GTGAGATTGG GATTATCTAC	780
CTCAACAATC AAAATAAAAA GGGGATTATG CAACGGGTTG AAAAGTTAGG TCTGGAGGTC	840
ATCGAATTGA TTCCTTTCCA TACCCATATT TATCTCTGTG AGGGTCATCC TTTAGCCCAG	900
40 AAAGAGGAAT TAGTCATGGA GGATTTAGCG GATTTACCAA CGGTTCGTTT CACTCAAGAG	960
AAAGACGAGT ACCTTTATTA TTCAGAGAAC TTTGTCGATA CCAGCGCTAC TCACAGATGT	1020
45 TTAATGTGAC AGACCGTGCC ACCTTGAATG GTATTTTGGG GCGGACGGAC GCCTATGCGA	1080
CAGGTTCTGG ATTTTtagat AGTGACAGTG TTAATGGCAT TACAGTTATT CGTCTCAAGG	1140
ATAACCTAGA TAACCGCATG GTCTATGTTA AACGTGAAGA AGTGGAGCTT AGTCAAGCTG	1200
50 GGACTCTCTT CGTAGAAGTC ATGCAAGAAT ATTTTGATCA AAAGAGGAAA TCATGAAAAA	1260
AAGAGCAATA GTGGCAGTCA TTGTACTGCT TTAATTGGG CTGGATCAGT TGGTCAAATC	1320
55 CTATATCGTC CAGCAGATTC CACTGGGTGA AGTGCGCTCC TGGATTCCCA ATTTCTTAG	1380
CTTGACCTAC CTGCAAAATC GAGGTGCAGC CTTTCTATC TTACAAGATC AGCAGCTGTT	1440
ATTGCTGTC ATTACTCTGG TTGTCGTGAT AGGTGCCATT TGGTATTTAC ATAAACACAT	1500
60 GGAGGACTCA TTCTGGATGG TCTTGGGTTT GACTCTAATA ATCGCGGGTG GTCCTGGAAA	1560



-43-

CTTTATTGAC AGGGTCAGTC AGGGCTTTGT TGTGGATATG TTCCACCTT

1609

## (2) INFORMATION FOR SEQ ID NO:4:

5

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 763 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

## (ii) MOLECULE TYPE: DNA (genomic)

15

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTAACCTAGG GCCCCAAGTC CATAACTTGC TTGACGCATG CTATCACTAA CAGATAAAAAG 60  
GGCTTCTTCT GTTGAGCGAA TAATGACTGG CAACACCATG ATGACTGAGG TTAAGATTCC 120  
TGATAACAGA GAGTATTGAA AACCTAAGAA GACTACAAAG AAGAGCATGC CAAACAGACC 180  
AAAAACAATG GAAGGAATCC CAGACAAGGT ATCTGAGGCC AATCGCATGA TTTTAACACA 240  
AAGGGAATCT TTTTTTGTAT ATTCCACAAG ATAAAAACCA GCAAAAATCC CTATGGGCAA 300  
GGCTAAAAGA AGAGCACCAA AGACCAGAAT AACGGTGGAA ATAATCGCTG GCATAAGGGA 360  
AATGTTCTCA GAAGTATAAG TCCAAGAAAA GAGGGATAGA CTTAGATGAG GTAAACCTTT 420  
GATGAGGATA AAACCAATGA TTAAAAAGAG AGAGCCAAAG GTTAAAGCTG AAAACAATA 480  
AACGAGAAGT TTTAGCAGGT ATTTACTCAT AAGATGATT TCCTTTCAAG TAGCCAAAGT 540  
AGGCATTAAT CAAGAGAATA AGGAAAAGA GAACGCTGA GGTGCAATA AGGGCTTCCC 600  
TATGCTGACC TGATGCGTAA GCCATTTCCA GAACAATATT GGTGTTAAG GTTCTGGTTC 660  
CTGAAAAGAG TCCACTTGGA ATAATCGGCT GGTGCTGC CACCAAATA ACTGCCATGG 720  
TTTCACCTAC TGC GCGACCG ATGCCTAAAA TAACTGCTGA AAA 763

50

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 897 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

55

## (ii) MOLECULE TYPE: DNA (genomic)

60

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

-44-

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGGTCTGTTT TGGCCTTGGC GGCTTCAGGT GGTTCAGGAG CTTGGCAGGG AGCTGGTCTC 60  
ATGTTGGTGT ATACGCTGGG CTTGGCGCTA CCATTCTTGC TTCTAGCTCT GACCTCTAGT 120  
10 TATGTTTTGA AACATTTCCG AAAACTTCAT CCCTATCTCG GAATCCTCAA AAAAGTGGGT 180  
GGTTTTCTCA TTATTGTGAT GGGATTCTTG GTTCTGTTTG GAAATGCTTC AATTTTAAGT 240  
15 CAATTATTTG AATAAAATGG AAAGGAATAT CAATATGAAA AAATGGCAA CATGTGTTCT 300  
TGGAGCAGGT TCGCTCCTTT GTTTGACGGC TTGTTCAGGC AAGTCCGTGA CTAGTGAACA 360  
CCAAACGAAA GATGAAATGA AGACGGAGCA GACAGCTAGT AAAACAAGCG CACTAAAAGG 420  
20 GAAAGAGGTG GCTGATTTTG AATTGATGGG AGTAGATGGC AAGACCTACC GTTTATCTGA 480  
TTACAAGGGC AAGAAAGTCT ATCTCAAATT CTGGGCTTCT TGGTGTTCCTA TCTGTCTGGC 540  
25 TAGTCTTCCA GATACGGATG AGATTGCTAA AGAAGCTGGT GATGACTATG TGGTCTTGAC 600  
AGTAGTGCA CCAGGACATA AGGGAGAGCA ATCTGAAGCG GACTTTAAGA ATTGGTATAA 660  
GGGATTGGAT TATAAAAATC TCCCAGTCCT AGTTGACCCA TCAGGCAAAC TTTTGGAAAC 720  
30 TTATGGTGTC CGTTCTTACC CAACCCAAGC CTTTATAGAC AAAGAAGGCA AGCTGGTCAA 780  
AACACATCCA GGATTCATGG AAAAAGATGC AATTTTGCAA ACTTTGAAGG AATTATCCTA 840  
35 GGAGGCGTCT TATGAATGAT AAGTTAAAAA TCTTCTTGTT GCTAGGAGTA TTTTTTC 897

## (2) INFORMATION FOR SEQ ID NO:6:

40 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3499 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

## 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTTCTTTTTC CTAGGTGATT TTAATGAGGT TGAAATTC AAATGTATTAG AATCATTTGG 60  
CTTTAAAGGT CGAAAAGGAG ATGTGAAGGT TCAGTATTGT CAACCTTATT CTAATATCCT 120  
60 TCAGGAAGGT ATGGTTCGGA AAAATGTGGG ACAATCCATT TTGGAATTAG GTTATCATT 180

-45-

	CCGTTCTAAA TATGGTGATG AGCAACATTT ACCCATGATT GTAATGAATG GTTACTTGG	240
	TGGATTGCT CACTCTAAGC TCTTTACAAA TGTCCGTGAA AATGCTGGAT TAGCTTATAC	300
5	CATTTCAAGT GAGCTTGATT TATTTAGTGG ATTCTTGAGG ATGTATGCTG GTATCAATCG	360
	AGAAAATCGT AACCAAGGCTC GTAAAATGAT GAATAATCAA CTGCTTGATT TAAAAAAGG	420
	TTATTTTACA GAGTTTGAGT TAAATCAGAC CAAGGAAATG ATTCGTTGGT CGTTGTTACT	480
10	TTCTCAAGAT AATCAATCTT CATTGATTGA ACGTGCTTAT CAAAATGCCT TATTTGAAA	540
	ATCTTCAGCA GACTTTAAAA GTTGGATTGC AAAGCTTGAA CAAATTGACA AAGATGCTAT	600
15	TTGTAGAGTA GCTAATAATG TGAAACTACA AGCGATTTC TTTATGGAAG GAATAGAATG	660
	ACAAAGGTTG TTTTGAAGA AAAATACTAT CCAGCTGTAA AAGAAAAGGT TTATCGAACT	720
	CGTTTGGCCA ACGGATTGAC AGTTGCTCTT TTGCCTAAAA AGGAATTTAA AGAGGTTTAC	780
20	GGGAGTGTCA CTGTACAGT TGGTTCGGTA GATACGTTT TCACAGAAGT TGACGGATAT	840
	GTAAACAAT ATCCTGGAGG AATTGCTCAT TTTCTGAAC ATAAATTATT TGAGAGAGAA	900
25	GATTCTAGTG ATTTGATGTC GGCTTTTACG AGTCTAGGTG CAGATAGTAA TGCCTTTACA	960
	AGCTTTACAA AAACAACTA TCTTTTTTCA GCAACGGATT ATTTTTTAGA AAATTTAGAT	1020
	TTACTTGATG AATTGGTAAC ATCAGCACAC TTTACTGAAG CTTCCATTCT GACAGAGCAG	1080
30	GATATTATTC AGCAAGAACG AGAAATGTAC CAAGATGATC CAGATTCGTG TTTATTCTTT	1140
	TCAACTTTAG CGAATTGTA TCCTGGTACA CCTTTAGCAA CTGATATAGT TGGAAGTGAG	1200
35	GAGTCCATT CCCAAATCAA TCTAACTAAT TTGCAAGAAA ATTTTACAAA GTTTTACAAA	1260
	CCTGTAAACA TGTCTCTGTT TTTAGTTGGT AATTTTGATG TGGAGCGAGT ACAGGACTAT	1320
	TTTGAAAGCA AAGAACTGAA AGATTCAGAT TTTAGGAAG TAGCAAGAGA AAAGTTGTTT	1380
40	TTACAGCCTG TAAAGCCAAC AGATAGTATG AGAATGGAAG TATCTTCTCC CAACTAGCG	1440
	ATTGGAGTTA GAGGTAAGCG AGAAGTTTCT GAAGCGGATT GCTATCGACA TCATATTTTA	1500
45	TTAAAATTAT TGTTTGCAAT GATGTTTGGT TGGACTTCGG GATCGTTTTT AAAAAATGTTA	1560
	TGAATCAGGT AAAATTGATG CGTCCTTATC TCTGGAAGTT AAATAACAAG TCGCTTTCAT	1620
	TTTGTCATGT TGACAAATAGA TACGAAAGAG CCAGTTGCTT TGTCTCATCA ATTTAGGAAG	1680
50	GCTATTCGTA ATTTTACAAA GGATTTAGAT ATTACAGAGG AACATTTAGA TATTATCAAA	1740
	AGAGAGATGT TTGGCGAATT TTTCAGTAGC ATGAACTCTC TTGAATTTAT TGCAACGCAA	1800
55	TATGATGCTT TTGAAAATGG TGAGACAATT TTTGATTTGC CGAAAATTTT ACAGGAAATT	1860
	ACTTTAGAGG ATGTCCTTGA TGCTGGACAT CATTTAATAG ATGATGGTGA CATAGTTGAT	1920
	TTTACAATAT TCCCATCGTA GTAACCTATC ATAATAGACA CTAGAAAGAA GGGATGACAA	1980
60	GTATGAGAAA AAAAACAATT GGAGAGGTTT TACGATTAGC TAGAATCAAT CAGGGATTGA	2040

-46-

5 GTTTAGATGA ATTGCAGAAA AAGACAGAAA TCCAGTTAGA TATGTTGGAA GCAATGGAAG 2100  
 CAGACGATTT CGATCAACTT CCAAGTCCTT TTTACACGCG TTCTTTCTTG AAAAAATATG 2160  
 CATGGGCTGT TGAGTTAGAT GACCAAATTG TTTTGGATGC TTATGATTCT GGGAGTATGA 2220  
 TTACTIONTATGA GGAAGTAGAT GTTGATGAAG ATGAGTTGAC AGGTCGTAGA CGTTCAAGTA 2280  
 10 AGAAAAAGAA GAAAAAACA TCATTTTTAC CTTTATTTTA TTTTATCCTT TTTGCTTTAT 2340  
 CGATTTTAAT TTTTGTGACT TATTATGTTT GGAACATAT TCAAACCTCAA CCAGAGGAGC 2400  
 CTTCTCTTTC TAATTACAGT GTGGTTCAAT CAACAAGTTC AACTAGCTCT GTTCCCCACT 2460  
 15 CCTCAAGTAG TAGTTCTTCT AGTATAGAAT CAGCTATAAG TGTATCAGGC GAAGGAAATC 2520  
 ATGTAGAAAT CGCTTATAAG ACAAGTAAGG AAACAGTTAA ATTGCAATTG GCAGTTTCAG 2580  
 20 ATGTTACAAG TTGGGTCAGT GTTTCAGAAA GCGAACTTGA GGGCGGTGTA ACCTTATCGC 2640  
 CAAAGAAGAA AAGTGCAGAA GCAACAGTTG CAACTAAAAG TCCTGTAACA ATTACGTTAG 2700  
 GTGTTGTAAG AGGTGTTGAT TTGACAGTAG ATAATCAGAC TGTGATTTA TCGAAATTAA 2760  
 25 CAGCTCAGAC TGGACAAATC ACTGTAACCT TTACTAAAAA TTAAGGAAAA ACGAATGAAA 2820  
 AAAGAACAAA TTCCAATCT CTTAACAATA GGTGGAATTC TCTTTATACC TATTTTTATC 2880  
 30 TTTATTTTAA CGATAGGAAA TTCGATAGAG AGTCATATAG TTGCAGCTAT TATCTTTGCT 2940  
 GTTGCCAGTA TTACCGACTA TTTAGATGGA TATTTAGCTC GTAAATGGAA TGTGGTCAGT 3000  
 AATTTTGGTA AATTTGCAGA TCCTATGGCG GATAAGTTAC TAGTTATGTC GGCTTTTATT 3060  
 35 ATGTTGATTG AGTTAGGTAT GGCTCCGGCT TGGATTGTTG CAGTGATTAT CTGTCGTGAG 3120  
 TTAGCTGTGA CAGGTTTAAG GCTTTTATTG GTTGAACTG GTGGAACAAT TTTAGCAGCA 3180  
 40 GCAATGCCTG GAAAAATTAA AACTTTTAGT CAGATGTTTG CTATTATTTT CTGCTATTA 3240  
 CATTGGACTT TGCTTGGTCA AGTTCTACTT TATGTAGCCT TATTTTTCAC TATCTACTCT 3300  
 GGCTATGACT ATTTCAAGGG TAGTGCCTAT GTATTTAAAG GGACATTTGG TTCGAAATGA 3360  
 45 AATCAATAAT TGATGTAAAA AATCTTCTT TTCGCTATAA AGAAAATCAG AACTACTACG 3420  
 ATGTGAAGGA TATTACGTTT CACGTGAAAC GTGGAGAATG GCTTTCGATT GTAGGGCATA 3480  
 50 ATGGTAGTGG TAAATCAAC 3499

## (2) INFORMATION FOR SEQ ID NO:7:

- 55 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 821 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 60 (ii) MOLECULE TYPE: DNA (genomic)

-47-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

10	ATTTTGAAT AATCAAGCGG AACCAAGAGG TCTTCGTCCT TCATCTTGTT AATCATGTAT	60
	TCACTTGGAA TGGCAATATC GTAGGTCGTT CCACCCTGCT TTATCTTAGT GTACATGGCT	120
	TCGTTGGAGT CAAAAGCCTC GTACTGAACT TGAATTCCTG TTTCTTCTGT AAACTGAGTC	180
15	AAGAGTTCAG GATCGATATA GTCTCCCCAG TTATAGATAA CCAATTTTGT ACTATCTCGA	240
	CTATTGATTT TACTATCTAA ATGAGTCGCA ATTCCCCACA AGACAAGGAT AATCGCTGCA	300
20	ATTCTGCTA AAATGAATAG ATTTTTTTCA TGCTTGCTCC TCCTTCTCAC GAGAGATAAA	360
	GTAATAACCT ACAACTAGGA TAATACTAAA GAGAAAGACT AGAGCAGAAA GGGCATTGAT	420
	TTCTAGCGAA ATCCCCTTGC GAGCACGAGA GTAAATCTCG ACTGATAGGG TTGAAAAGCC	480
25	ATTTCTGTT ACAAGAAGG TCACGGCAAA GTCATCTAAC GAATAGGTGA AGGCCATGAA	540
	ATAACCAGCA ATGATAGACG GAGTCAGGTA AGGAAGCATG ATTTCTTAA ACATCTGAAA	600
30	TTGACTAGCT CCCAAGTCAT AGGCCGCATG AATCATGTCG CCATTCATTT CCTTGAGTCG	660
	GAGGCAAGAC CATCAAGACC ACGATAGGAA TGGAGAAGGC CACATGACTA GATAGAACGG	720
	TCAAAAAGCC AAGTGAAAAC TTGAGTTGGG TAAAGAGAAT CAAGAAGTAG CACCAATCAT	780
35	AACGTCAGGC GCAACCATGA GGATATTATT GAGTGATAGA A	821

(2) INFORMATION FOR SEQ ID NO:8:

40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1309 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

	GGCCGGTGCC ACAGTCCAAG CTATCGGTAT CGTGATTGAG AAATCCTTCC AAGATGGTCG	60
60	TGATTTGCTT GAAAAAGCAG GCTACCCTGT CCTATCACTT GCTCGCTTGG ATCGTTTTGA	120
	AAATGGTCAG GTCGTATTTA AGGAGGCAGA TCTCTAATGC AAATCAAGA AAAACACTCG	180

-48-

CAAGCAGCCG TTCTCGGCTT GCAGCACTTA CTAGCCATGT ACTCAGGATC TATCCTGGTT 240  
CCCATCATGA TTGCGACAGC CCTTGGCTAT TCAGCTGAGC AGTTGACCTA CCTGATTTC 300  
5 ACAGATATCT TCATGTGTGG GGTGGCAACC TTCCTCCAAC TCCAACCTCA CAAATACTTT 360  
GGGATTGGAC TCCCAGTCGT TCTTGGAGTT GCATTCCAGT CGGTCGCTCC CTTGATTATG 420  
10 ATTGGGCAAA GCCATGGTAG TGGCGCTATG TTTGGTGCCC TTATCGCATC TGGGATTTAC 480  
GTGGTTCTTG TTTCAGGCAT CTTCTCAAAA GTAGCCAATC TCTTCCCATC TATCGTAACA 540  
GGATCTGTTA TTACCACGAT TGGTTTAACC TTGATCCCTG TCGCTATTGG AAATATGGGA 600  
15 AATAACGTTC CAGAGCCAAC TGGTCAAAGT CTCTTGCTTG CAGCTATTAC TGTTCTGATT 660  
ATCCTCTTGA TCAACATCTT TACCAAAGGA TTTATCAAGT CTATCTCTAT TTTGATTGGT 720  
20 CTGGTTGTTG GAACTGCCAT TGCTGCTACT ATGGGCTTGG TGGACTTCTC TCCTGTTGCG 780  
GTAGTCCACT TGTCCATGTC CCAACTCCAC TCTACTTTGG GATGCCAACC TTTGAAATCT 840  
CATCTATTGT CATGATGTGT ATCATCGCAA CGGTGTCTAT GGTGAGTCA ACTGGTGTTC 900  
25 ATCTAGCCTT GTCTGATATC ACAAAGATC CAATCGACAG CACGCGCCTT CGCAACGGTT 960  
ACCGCGCAGA AGGTTTGGCG GTACTTCTCG GAGGAATCTT TAACACCTTC CTTACACCG 1020  
30 GATTTTCACA AAACGTTGGT TTGGTTAAAT TGTCAGGCAT CAAAAACGC CTGCCAATCT 1080  
ACTACGCAGG TGGTTTCCTG GTTCTCCTTG GACTGCTTCC TAAGTTTGGT GCCCTTGCCC 1140  
AAATCATTCC AAGCTCCGTC CTCGGCGGTG CCATGCTGGT GATGTTTGGT TTTGTATCTA 1200  
35 TTCAAGGGAT GCAATCCTC GCCCGAGTTG ACTTTGTAAC AATGAACACA ACTTCCTTAT 1260  
CGCAGTGTTT CAATCGCTGC AGGTGTCGGT CTCAACAACA AGTAATCTC 1309

## 40 (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 1031 base pairs  
    (B) TYPE: nucleic acid  
45     (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

50 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

60 TTAAAGTTCC AGTTTATCTA GGTCTTCAT TTGCCTTTAT CACAGCTATG TCACTGGCTA 60  
TGAAAGAAAT GGGGGGTGAT GTATCTGCTG CCCAACAGG GGTATCTTG ACTGGTTTGG 120

-49-

5 TCTATGTCCT TGTGCTACC AGCATCCGAT TTGTAGGAAC AAAATGGATT GATAAACTCT 180  
TGCCACCAAT CATTATCGGT CCTATGATCA TCGTTATCGG TCTTGGACTT GCAGGTTGAG 240  
CTGTTACCAA TGCAGGTCTT GTAGCAGACG GAAATTGGAA AAATGCTCTG GTAGCCGTTG 300  
TTACTTTCCT AATTGCTGCC TTTATCAATA CAAAAGGAAA AGGCTTCCTA CGAATCATTC 360  
10 CATTCTCTT TGCCATTATC GGTGGTTACC TTTTCGCACT AACTCTTGGC TTGGTTGACT 420  
TTACACCAGT TCTTAAAGCC AACTGGTTTCG AAATTCCTGG TTTCTACTTG CCATTTAGCA 480  
CAGGTGGTGC CTTTAAAGAG TACAATCTTT ACTTTGGTCC AGAAGCCATC GCTATCTTGC 540  
15 CAATCGCTAT CGTAACAATT TCTGAACATA TCGGAGACCA TACTGTTTGG GGTCAAATCT 600  
GTGGCCGTC ATTCTTAAAA GAACCAGGTC TTCATCGTAC TCTTCTTGGT GACGGTATCG 660  
20 CAACTTCTGT TTCTGCCTTC CTTGGTGGAC CAGCCAATAC AACTTACGGA GAAAATACAG 720  
GGGTTATCGG TATGACTCGT ATCGCTTCTG TCTCAGTTAT CCGTAACGCT GCCTTCATCG 780  
CGATTGCCCT CAGCTTCCTT GGTAAATTCA CTGCCTTGAT TTCAACTATT CCAAACGCTG 840  
25 TACTTGGTGG TATGTCAATC CTCTCTATG GGGTTATCGC CAGCAATGGT TTGAAAGTCT 900  
TGATTAAAGA ACGTGTGAT TTCGCTCAAA TCGGAAACCT CATCATCGCA AGTGCTATGT 960  
30 TGGTTCTTGG ACTTGGAGGA GCTATCCTTA AACTTGGTCC AGTACACTTT CAGGTACTGC 1020  
CCTTTCAGCC A 1031

## (2) INFORMATION FOR SEQ ID NO:10:

- 35 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 568 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
40 (D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
45 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ACAGTTTAAT CATTGCCTTG GCTACAACCC TCATTGCGAT TATTATTCTT GCTATGGCAG 60  
55 CCTATGGTAT TGTTCGATTC TTTCTTAAAT TGGGAGCAAT CATGTCGAGA CTACTCGTCA 120  
TTACCTACAT TTTCCACCA ATTTTGTAG CAATTCCTTA TTCAATTGCC ATTGCTAAAG 180  
TTGGGTTAAC AAATAGTTTA TTTGGCTTGA TGATGGTTTA TCTATCTTTT AGTGTTCAT 240  
60 ATGCAGTTTG GCTCTTAGTT GGATTTTTC CAAACAGTTCC AATTGGAATT GAAGAAGCGG 300

-50-

CTAGAATTGA TGGTGCAAAT AAATTTGTTA CGTTTATAA AGTTGTGCTA CCGATTGTAG 360  
CACCAGGTAT TGTAGCAACA GCTATTTATA CATTTATCAA TGCTTGAAT GAATTTCTGT 420  
5 ATGCCTTGAT TTTGATTAAAC AATACAGGAA AGATGACAGT AGCAGTAGCC CTCGTTTAC 480  
TTAATGGTTC AGAAATACTA GACTGGGGAG ATATGATGGC AGCGTCTGTT ATTGTAGTTC 540  
10 TTCCATCAAT TATTTCTTCT CTATCACC 568

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
15 (A) LENGTH: 468 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
20 (ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
(iv) ANTI-SENSE: NO  
25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

30 ACCAAAGACT TAGCTTCTTC AAAAAGCGGA TCACCACCAG CATCTCCATC CGAAAATTCT 60  
CCTTCATTTT CAGAAACCTC ACCTGGATCA AAACCTCAT CGTAGTCTGC ATCTGCCTGA 120  
35 GTCTTGATGA AGTTCACAAT GCGCTCAACA TCGTCATCCG AGATAAAGGA GCCTTGGAGA 180  
CGAACTGGAT GATTTTTCATC AATCGGTTTA AAGAGCATGT CTCCTCGACC AAGAAGTTTT 240  
TCTGCTCCAT TTTCATCCAA AATCGTACGG GAGTCTGTTC CTGATGAAAC CGCAAATGCT 300  
40 ACACGAGATG GAACATTGGC CTTAATCAAA CCAGAGATGA CATCAACAGA TGGACGCTGA 360  
GTTGCAAGAA TCATGTGGAT ACCTGCAGCA CGCGCCTTCT GCCCAAGACG GATGATAGCA 420  
45 TCTTCCACTT CTTGCTGGC CACCATCATG AGGTCAGCCA ACTCATCC 468

## (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:  
50 (A) LENGTH: 466 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
55 (ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
(iv) ANTI-SENSE: NO  
60



-51-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

5 AAGCTGACAA TCTTTTCTGC AGTTGGAGCA TCCCAGAAGG ATACACCACT AAGGATGCGA 60  
CCTGCCTTGC TATCAACAAT AATGTCTTGA ACCTTGTAAGT CATCTCCATA GACCAAGAAC 120  
CATTGCTTGG TACAATCTTC ACGATAAACA CTAAAATAAG TCGAACGAGT CAAATCATTG 180  
10 CGGAACATAT TTTTAAAGAG ATAGTTATCT GCATCAATAA CATAGCTGTT GGCCAATTCT 240  
TCTTTTACAA GATAGAGAGA GTAAAAGTTA TTGTAGTCAG CGTATTTATC ATTGAAAACG 300  
15 AGACGAACAC CGTATTTCTC TTTCAAGTAA TCGAATTGTT CTTTAAGATA ACCAACAATG 360  
ATGATGATGT CATTGATTCC TTTTCTTTG AGAACTCAA TTTGGTACTC AATCAAAGGT 420  
TTTTGATTAA CCTGAACCAA GGCTTTAGGG GTATTTTCAG TCATAG 466

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 1040 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

30

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

40 CACATAATCT GTATATTGAC TATAAGTTTT AAAAACAAT TTTTAAGCTC TTCCTTGTCT 60  
TCTCTAACCA AGCGTGTTAT AATGAATACT GCTCAAGCGA CCTTCAATCG TGAAGCACAC 120  
ACGACCTTCA ATCGTGAATA AACGAATAGA TGGGAGACTT ACCATGAGTG ATAACTCTAA 180  
45 AACACGTGTT GTCGTGGGGA TGAGTGGTGG TGTGATTCTG TCGGTGACGG CTCTTTTGCT 240  
CAAGGAGCAG GGCTACGATG TGATCGGTAT CTTTATGAAG AACTGGGATG ACACAGATGA 300  
50 AAACGGCGTC TGTACGGCGA CCGAAGATTA CAAGGATGTG GTTGCGGTGG CAGATCAGAT 360  
TGGCATTCCC TACTACTCTG TCAATTTTGA AAAAGAGTAC TGGGACCGCG TTTTGTAGTA 420  
TTTCTAGCT GAATACCGTG CAGGGCGCAC GCCAAATCCG GACGTTATGT GCAACAAGGA 480  
55 AATCAAGTTC AAGGCCTTTT TGGACTATGC CATGACCTTG GGGGCAGACT ATGTAGCGAC 540  
TGGGCACTAT GCTCGAGTGG CGCGTGATGA GGATGGCACT GTTCACATGC TTCGTGGCGT 600  
60 GGACAATGGC AAGGATCAGA CCTATTTTCT CAGCCAATT TCGCAAGAAC AACTTCAAAA 660

-52-

AACCATGTTC CCACTAGGAC ATTTGAAAAA GCCTGAAGTT CGAAAACTAG CAGAAGAAGC 720  
AGGTCTTTTCG ACTGCTAAGA AGAAAGACTC GACAGGGATT TGCTTTATCG GAGAAAAGAA 780  
5 CTTTAAAAAC TTTCTCAGCA ACTACCTGCC AGCTCAGCCT GGTCGTATGA TGACTGTGGA 840  
TGTCGTGAT ATGGGCGAGC ATGCTGGTCT TATGTACTAT ACAATCGGTC AGCGTGGCGG 900  
ACTCGGTATC GGTGGGCAAC ACGGTGGTGA CAATGCCCCCT TGGTTCGTTG TCGGAAAAGA 960  
10 TCTAAGCAAG AATATTCTCT ATGTAGGCCA AGGTTTCTAC CATGATTCGC TCATGTCAAC 1020  
CACTAGAGGC TAGCCAAGTC 1040

## 15 (2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3071 base pairs  
(B) TYPE: nucleic acid  
20 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCGGGGATCT GATAGCCAAT AGAAAACCGC AGAGTCAAAG GGTTTTGTAT GAATTGCGAG 60  
35 ATCGTTTGAA GAGAAATCAG TTTATACTCA ATGATACCAA TCCGGATATT GTCATTTCCA 120  
TTGGCGGGGA TGGTATGCTC TTGTCGGCCT TTCATAAGTA CGAAAATCAG CTTGACAAGG 180  
40 TCCGCTTTAT CGGTCTTCAT ACTGGACATT TGGGCTTCTA TACAGATTAT CGTGATTTTG 240  
AGTTGGACAA GCTAGTGACT AATTGCAAC TAGATACTGG GGCAAGGGTT TCTTACCCTG 300  
TTCTGAATGT GAAGGTCTTT CTTGAAAATG GTGAAGTTAA GATTTTCAGA GCACTCAACG 360  
45 AAGCCAGCAT CCGCAGTCTG ATCGAACCAT GGTGGCAGAT ATTGTAATAA ATGGTGTTC 420  
CTTTGAACGT TTTCGTGGAG ACGGGCTAAC AGTTTCGACA CCGACTGGTA GTACTGCCTA 480  
50 TAACAAGTCT CTTGGCGGTG CTGTTTACCA CCCTACCATT GAAGCTTTGC AATTAACGGA 540  
GATTGCCAGC CTTAATAATC GTGTCTATCG AACATTGGGC TCTTCCATTA TTGTGCCTAA 600  
GAAGGATAAG ATTGAACTTA TTCCAACAAG AAACGATTAT CATACTATTT CGGTTGACAA 660  
55 TAGCGTTTAT TCTTTCCGTA ATATTGAGCG TATTGAGTAT CAAATCGACC ATCATAAGAT 720  
TCACTTTGTC GCGACTCCTA GCCATACCAG TTTCTGGAAC CGTGTTAAGG ATGCCTTTAT 780  
60 CGGTGAGGTG GATGAATGAG GTTTGAATTT ATCGCAGATG AACATGTCAA GGTTAAGACC 840

-53-

	TTTTTAAAAA AGCACGAGGT TTCTAAGGGA TTGCTGGCCA AGATTAAAGTT TCGAGGTGGA	900
	GCTATTCTGG TCAATAATCA ACCGCAAAAT GCAACGTATC TATTGGACGT TGGAGACTAC	960
5	GTTACCATTG ACATTCCCGC TGAGAAAGGC TTTGAAACCT TGGAGGCTAT TGAGCTTCCA	1020
	TTAGATATTC TCTATGAGGA TGACCACTTT CTAGTCTTGA ATAAACCCTA TGGAGTGGCT	1080
10	TCTATTCTTA GTGTTAATCA CTCTAATACC ATTGCCAATT TTATCAAGGG TTAATATGTC	1140
	AAGCAAAATT ATGAAAATCA GCAGGTTTAC ATTGTTACCA GACTAGATAG GGACACTTCT	1200
	GGCTTGATGC TCTTTGCCAA GCACGGTTAT GCCCATGCAC GATTAGACAA GCAGTTGCAG	1260
15	AAGAAATCTA TCGAGAAACG CTAATTTGCT TTGGTTAAGG GAGATGGACA TTTGGAGCCA	1320
	GAAGGGGAAA TTATTGCTCC GATTGCGCGT GATGAAGATT CCATTATTAC CAGACGAGTG	1380
20	GCTAAAGGCG GAAAGTATGC CCATACTTCA TACAAGATTG TAGCTTCTTA TGGAAATATT	1440
	CACTTGGTCT ATATTCACCT GCACACTGGT CGAACCCATC AAATCCGAGT CCATTTTTCT	1500
	CATATCGGTT TTCCTTTGCT GGGAGATGAT TTGTATGGTG GTAGTCTGGA AGATGGTATT	1560
25	CAACGTCAGG CTCTGCATTG CCATTACCTA TCCTTTTATC ATCCATTTTT AGAGCAAGAC	1620
	TTGCAGTTAG AAAGTCCCTT GCCGGATGAT TTCAGTAACC TTATTACCCA GTTATCAACT	1680
30	AATACTCTAT AAAAATGTC TCAGAGTATA ATTATTATCT TAAAGGAGAA AACTCATGGA	1740
	AGTTTTTGAA AGTCTCAAAG CCAACCTTGT TGGTAAAAAT GCTCGTATCG TTCTCCCTGA	1800
	AGGGGAAGAG CCTCGTATTC TTCAAGCAAC AAAACGCTTA GTAAAAGAAA CAGAAGTGAT	1860
35	TCCTGTTTTG CTGGGAAATC CTGAAAAAAT TAAATTTAT CTTGAAATTG AAGGAATCAT	1920
	GGATGGTTAT GAGGTATCG ACCCTCAACA TTATCCTCAA TTTGAAGAAA TGGTTTCTGC	1980
40	CTTGGTGGAG CGTCGCAAGG GCAAAATGAC TGAAGAAGAT GTACGCAAGG TTTTGGTTGA	2040
	AGATGTCAAC TACTTTGGTG TGATGTTGGT TTAATTTGGC TTGGTTGATG GAATGGTGTC	2100
	AGGAGCGATT CACTCAACAG CTTCAACAGT TCGCCAGCT CTACAAATCA TCAAACTCG	2160
45	TCCAAATGTA ACTCGTACTT CAGGAGCCTT CCTCATGGTT CGTGGTACGG AACGTTACCT	2220
	ATTTGGAGAC TGTGCCATTA ATATCAATCC AGATGCAGAA GCCTTGGCTG AAATTGCCAT	2280
50	CAACTCAGCA ATCACAGCTA AGATGTTTGG CATCGAACCT AAAATTGCCA TGTGAGCTA	2340
	TTCTACTAAA GGTTCAGGGT TTGGTGAAAG CGTTGATAAG GTCGTTGAAG CAACTAAAT	2400
	TGCTCACGAC TTGCGTCTG ACCTTGAAAT CGATGGTGAG TTGCAATTG ATGCGGCCTT	2460
55	TGTTCCCGAA ACTGCAGCTC TGAAAGCTCC GGGAGTACA GTAGCTGGTC AAGCAAATGT	2520
	CTTCATCTTC CCAGGTATCG AGGCAGGAAA TATCGGTTAC AAGATGGCTG AACGCCTGGG	2580
60	TGGCTTTGCG GCTGTAGGAC CTGTTTGTCA AGGTTTAAAC AAGCCAGTTA ATGATCTTTC	2640
	TCGTGGATGT AATGCAGATG ATGTTTACAA GTTGACCCTC ATCACAGCAG CTCAAGCAGT	2700

-54-

TCATCAATAG TGAAACTAT AAAGTGATAT ACTATGCTAT ACTGTAGTTA TGAAACTATG 2760  
TACGAAAAGC ACTGCCATTA ATTCCTGAGA ACTAAATTAC TGATTGGTGT CAAAAGGAA 2820  
5 AACTTCCAAG CGATGATATC CTGTCTATAC ACGACCTATA GAAATCTGTA ATATACATGT 2880  
CCGTAAAACG ATAAATTCCC TTTTGTATTT TAAATGAGTA TGAAAAGAGA ATTTTCCGGC 2940  
10 TCTTTGTCAA CTGTAGTGGG TTGAAAAAAA GCTAAGCTCG AGAAAGGACA AATTTTGTCC 3000  
TTTCTTTTTT GATATTCAGA GCGATAAAAA TCCGTTTTTT GAAGTTTCA AAGTTTCGAC 3060  
TCTAGAGGAT C 3071

15 (2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:  
20 (A) LENGTH: 720 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

35 TTTCCATGGT ATGGTAAAGG TTTTCTTTT TTTTAAAAGG AAAACGAGAA GAGGAGGTTG 60  
TTATGAAAGC AAGCATTGCC TTGCAAGTTT TACCCCTAGC ACAGGGGATT GATCGGATAG 120  
CTGTTATTGA TCAGGTCATT GCTTATCTGC AAACCTAAGA AGTGACGATG GTAGTGACAC 180  
40 CATTTGAAAC GGTCTTGGA GGGGAGTTT ATGAGCTTAT GCGCATTCTA AAAGAAGCGC 240  
TGGAAGTGGC AGGGCAGGAG GCAGACAATG TCTTTGCCAA TGTCAAATA AATGTAGGAG 300  
45 AGATTTTAAG TATTGATGAG AAACCTGAAA AGTATACTGA GACGACACAT TAGTCTATTG 360  
GGCTTTCTCG GAGTATTGTC AATCTGGCAG TTAGCAGGTT TTCTTAACT TCTCCCCAAG 420  
TTTATCCTGC CGACACCTCT TGAAATTCTC CAGCCCTTTG TTCGTGACAG AGAATTTCTC 480  
50 TGGCACCATA GCTGGGCGAC CTTGAGAGTG GCTTTACTGG GGCTGATTTT GGGAGTTTGT 540  
ATTGCCTGTC TTATGGCTGT GCTCATGGAT AGTTTGACTT GGCTCAATGA CCTGATTTAC 600  
55 CCTATGATGG TGGTCATTCA GACCATTCCG ACCATTGCCA TAGCTCCTAT CCTGGTCTTG 660  
TGGCTGGGTT ATGGGATTTT TGCCCAAGAT TGTCTTGATT ATCTTAACAA CAACCTTTCC 720

60 (2) INFORMATION FOR SEQ ID NO:16:

-55-

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 852 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
- |     |  |
|-----|--|
| 60  | GCCGTCATAA TCATGCGCCG AATCCGTCCC CATTAAAATC TGGGTCTGTA AAGACAATGA  |
| 120 | CTCCATGACG TTGGTGTAGA CGCTGAATCC GCTCTATGTC CTGGTCATTG ATGGCAGAAC  |
| 180 | CTCGAGTCTC ATAGGTCTCC ACATCGAAAT AACGCTTGAG ATTGACCGTA TCATCACGAC  |
| 240 | CTTCAACCAC GATAACTTGG GAAATTCTCT CTTTCATTAC TTGCTGTCCA ATCCCAAAAA  |
| 300 | TGCGTTCTGC ATTTGCAGTC GTTGCTACCG CCAGCTCTTC TGTCGTCATA CCACGCAAGT  |
| 360 | CAGCGATAAA GTCGACCACA TAGCGAGTAT AGGCTGTTTT ATTTTCACGA CCACGCTTGG  |
| 420 | GTACAGGTGC TAAGTAAGGC GCATCTGTTT CTACCAACAT CTTGTCCAAA GGTAACCTCTT |
| 480 | TAGCTGCTTC TTGGAGGTCA GTTGCCCTTCT TGAAGGTCAC CACTCCTGAG AAGGAAATGG |
| 540 | TCATACCAAG ATCCCGGTAC CGAGCCCACT CAAGCGTCCC TGAAAATGAA TGCATGATAC  |
| 600 | CACCACGAGG ACCAACGCCC TCACTCTTGA TAATCTCATA GGTATCTTCC AGCGCATCAC  |
| 660 | GGGTATGGAC AACAAAAGGC AAATCCAAGT CCTTAGATAG CTGAATCTGA CGGCGAAAAA  |
| 720 | CCTGCTCCTG CACCTCTTGG GCGCTGTCAT CCAATGGTAG TCTAAGCCAA TTTCACCTAA  |
| 780 | AGCCACAACC TTGGAATGTT TTAAGTTATC CAACAAGTAA GCCTCAACTT CCTCTGTATA  |
| 840 | AGTACCAGCT TCTGTAGGAT GCCAACCAAT AGTCGCATAG AGCTGCTCAT ACTCATCTAC  |
| 852 | CAAACTCCAA GG  |
- (2) INFORMATION FOR SEQ ID NO:17:
- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 868 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

-56-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

5 GGGGATCCTC TAGAGTCGAT ATCTACGGTC TCAACCGTAC AGGACTGTTG AACGATGTAC 60  
TGCAAGTTCT TTCAAATACA ACCAAGAATA TTTCAACGGT CAATGCCCAA CCAACCAAGG 120  
10 ATATGAAGTT TGCTAATATC CATGTGTCCT TCGGTATTGC CAACCTCTCT ACACTGACCA 180  
CGGTTGTCGA TAAAATTAAG AGTGTGCCAG AAGTTTACTC TGTCAAACGG ACCAACGGCT 240  
AGTTGTCCTA GCTCTTACTA GAAAGGCTAT TATGAAAATC ATTATCCAAC GGGTTAAAAA 300  
15 AGCCCAAGTG AGTATAGAAG GCCAGATTCA GGGAAAAATC AATCAGGGAC TTTTATTGCT 360  
GGTTGGTGTT GGACCAGAGG ACCAAGAGGA AGATTGGAC TATGCTGTGA GAAAACTGGT 420  
20 CAATATGCGG ATTTTTCAG ACGCAGAAGG CAAGATGAAC CTGTCTGTCA AAGATATTGA 480  
AGGAGAAATC CTCTCTATTT CTCAGTTTAC CCTCTTTGCG GATACTAAGA AAGGCAATCG 540  
TCCAGCCTTT ACAGGGGCAG CTAAACCTGA TATGGCATCA GACTTCTATG ATGCTTTCAA 600  
25 TCAAAAATTA GCGCAAGAAG TGCCCGTTCA GACAGGTATC TTTGGAGCAG ATATGCAGGT 660  
TGAGCTGGTT AATAACGGAC CTGTTACCAT TATCCTAGAT ACTAAAAAGA GATAAGAAAG 720  
30 ACCAAGCCCA GTCGGCTTGG TCTTTCTCAT CGATCATAAA AATACTCCAA AAAGAAATCG 780  
GTTCTTGATA TGCTTGGGGG ACTCTTTTCA GGCTTTGGCA GATGCGATAG GAAGGGATGA 840  
GATGTCCTAG GGTGAGGAGA GTTCCCTG 868

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 1399 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

45.

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

50

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

55 CGGTCCTCGT CCGATTGACT CACACCTTAA GGC GTTTGAA GCTATGGGTG CCACTGCTAG 60  
CTACGAGGGA GATAACATGA AGTTATCTGC TAAAGATACA GGACTTCATG GTGCAAGTAT 120  
TTACATGGAT ACGGTTAGTG TGGGAGCAAC GATTAATACG ATGATTGCTG CGGTAAAGC 180  
60 AAATGGTCGT ACTATTATTG AAAATGCAGC CCGTGAACCT GAGATTATTG ATGTAGCTAC 240

-57-

TCTCTTGAAT AATATGGGTG CCCATATCCG TGGGGCAGGA ACTAATATCA TCATTATTGA 300  
 TGGTGTGAA AGATTACATG GGACACGTCA TCAGGTGATT CCAGACCGCA TTGAAGCTGG 360  
 5 AACATATATA TCTTTAGCTG CTGCAGTTGG TAAAGGAATT CGTATAAATA ATGTTCTTTA 420  
 CGAACACCTG GAAGGGTTTG TTGCTAAGTT GGAAGAAATG GGAGTGAGAA TGAAGTATC 480  
 10 TGAAGACAGC ATTTTGTGCG AGGAACAGTC TAATTTGAAA GCAATCAATA TTAAGACAGC 540  
 TCCTTACCCA GGCTTTGCAA CTGATTTGCA ACAACCGCTT ACCCCTCTTT TACTAAGAGC 600  
 GAATGGTCGT GGTACAATTG TCGAGTCGAT ACGATTACG AAAACGTGT AAATCATGTT 660  
 15 TTTGAACTAG CAAAGATGGA TCGGATATT TCGACAACAA ATGGTCATAT TTTGTACAGC 720  
 GGTGGACGTG ATTTACGTGG GGCCAGTGTT AAAGCGACCG ACTTAAGAGC TGGGGCTGCA 780  
 20 CTAGTCATTG CTGGGCTTAT GGCTGAAGGC AAAACTGAAA TTACCAATAT CGAGTTTATC 840  
 TTACGTGGTT ATTCTGATAT TATCGAAAAA TTACGTAATT TAGGAGCGGA TATTAGACTT 900  
 GTTGAGGATT AAACCGTAGA GGTGTTTATG AATATTTGGA CCAAATTAGC AATGTTTCT 960  
 25 TTTTTTGAAA CGGATCGCTT GTATTTGCGT CCTTCTTTT TTAGTGATAG TCAGGACTTC 1020  
 CGCGAGATAG CTTCAAATCC AGAAAATCTT CAATTTATTT TCCCAACGCA GGCAAGTCTG 1080  
 30 GAAGAAAGTC AATATGCACT GGCCAATTAC TTTATGAAGT CCCCTTTGGG AGTGTGGGCA 1140  
 ATTTGTGACC AGAAAAATCA ACAAATGATT GGTTCATTA AATTTGAGAA GTTAGATGAA 1200  
 ATCAAAAAAG AAGCTGAGCT TGGCTATTTT TTGAGAAAAG ATGCTTGGTC GCAAGGATTT 1260  
 35 ATGACAGAGG TTGTTAGAAA AATTTGTCAG CTTTCTTTT AGGAATTTGG CTTAAAACAA 1320  
 TTATCTATCA TTACCCACCT TGAAAATGAA GCTAGCCAAA GAGTTGCTCT TAAGTCTGGA 1380  
 40 TTTAGTTTGT TCCGTCAGT 1399

## (2) INFORMATION FOR SEQ ID NO:19:

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1779 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

55 (iv) ANTI-SENSE: NO

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AGATTGCTCT TGAACACGAT GAAATACCAA TTGGTTGTGT GATTGTCAAA GATGGGAAAA 60

-58-

	TCATTGGTCG TGGGCATAAT GCGCGTGAGG AATTACAGCG ACGGTTATGC ATGCGGAAAT	120
	TATGGCTATA GAGGATGCCA ACTTGAGTGC AGGAGACTGG CGCTTGCTGG ATTGCACACT	180
5	TTTTGTGACC ATTGAACCAT GTGTCATGTG TAGTGGGGCG ATTGGGCTTG CCCGTATTCC	240
	AAATGTGGTC TATGGGGCTA AAAACCAGAA ATTTGGCGCT GCTGGAAGTT TGTACGATAT	300
10	CTTGACAGAT GAGCGTCTTA ACCATCGTGT AGAGGTTGAA ACGGGAATTT TGAAGATGA	360
	ATGTGCAGCT ATCATGCAGG ACTTTTTTAG AAATAGACGG AAAAAATAAT TTTGCTTTTA	420
	AAATGAATAG GAATGTGATA TAATAAATAG TGGAGCAACA GTTCTGCGTG AAGCGGGTCA	480
15	GGGGAGGAAT CCCAGCAGCC CTAAGCGATT TGAATTGTGT GCTCTTTTTT TCGTGCTTTT	540
	TTCCGAATAA ATAAGATAGA ATAATCTAGA ATAAATGATA ATAGAAAAGA GAAAATTATG	600
20	AAAATTCGTG GTTTTGAATT GGTTCGAGT TTTACAGATG AAAATTTATT GCCCAAGCGT	660
	GAGACAGCGC ATGCGGCTGG TTACGACTTA AAGTTGCTG TCGTACAGT TGTTGCGCCA	720
	GGAGAGATTG TCTTGTTCC GACAGGGGTT AAGGCTTATA TGCAGCCGAC TGAGGTTCTC	780
25	TACCTCTATG ATCGTCTTC AAATCCTCGT AAGAAGGGCT TGGTTTTAAT TAACTCAGTT	840
	GGGGTCATTG ATGGGGATTA TTATGGAAAT CCTGGAAATG AAGGGCATAT TTTTGCGCAG	900
30	ATGAAGAATA TCACAGACCA AGAGTTGTT CTTGAAGTTG GGGAGCGTAT TGTCCAGGCT	960
	GTTCCTGCTA CTTTCTTAAT TGCAGATGGA GATGCAGCTG ATGGCGTTCG AACTGGTGGA	1020
	TTTGGATCGA CAGGGCACTA GAATGAAGAT TATCTTTGTA CGTCATGGGG AGCCAGATTA	1080
35	CCGTGAGTTA GAGGAGCGTT CTTATATAGG ATTTGGGATA GATTTGGCAC CCTTGTCTGA	1140
	GATGGGACGG CAGCAAGTCC AGAAATTGAG CAAAAATCCT TTAATCTCGT CAGCTGAAAT	1200
40	AATCGTATCT TCTGCAGTCA CAAGAGCTTT AGAAACGGCT TCGTATGTGG TCTGTGCTAC	1260
	GGGTCTTCCT TTAAGAGTAG AGCCTTTATT ACATGAATGG CAGGTCTATA AAACAGGAAT	1320
	AGAAAACTTT GAAACAGCTA GAAGACTGTT TTTAGAAAAC AAGGGGGAGT TGCTTCCTAA	1380
45	TAGTCCTATT CAATATGAGA CAGCTACGGA AATGAAGTCT CGGTTTCTAG AATGTATGTC	1440
	TAAGTATCGA GAACATCAGA CTGTGGTAGT GTTGCTCAT CGACTCTAGA GGAGCCAGTT	1500
50	TGTGCCAAAT GAGAAGATTG ATTTTGGCCA AGTGATTGAG TGTGAGTTAG AGATATAGAA	1560
	AGAGGTTTGT CATCGCAAAG AAAAAAGCGA CATTTGTATG TCAAAATTGT GGGTATAATT	1620
	CCCCATAATA TCTGGGACGT TGCCCCAACT GTGGGTCTTG GTCTTCTTTT GTGGAAGAGG	1680
55	TTGAGGTTGC CGAAGTTAAG AATGCGCGTG TGTCTTGAC AGGTGAGAAA ACCAAGCCCA	1740
	TGAAACTAGC TGAGGTGACT TCCATCAATG TCAATCGAC	1779
60	(2) INFORMATION FOR SEQ ID NO:20:	



-59-

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 3725 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

5	GCGGATCCTC TAGAGTCGAA AGATTACGAA GGTAAGAACC CTCTTTATTA CTGGTCACAT	60
20	CATGGTACAA CAAGCTGCAA CAGTATCTCT TATGGTTCTA TTCTTAGTAC CACAATTGCG	120
	CAATGCTTAC GGTACAGCAG CGATTGGTAT CATCTGTGGA CTTTACTGGG CAGTTAGTTC	180
25	AAATATGACT GTTGAGGCAA CTCAACGCTT GACTGGTGGT GCGGATTTG CGATTGGTCA	240
	CCAACAGCAA TTTGCAATCT GGTGTGTAGA TAAAGTAGCA GGACGCTTG GTAAGAAAGA	300
	AGAAAGTTTA GACAATCTTA AATTACCTAA GTTCCTCTCA ATCTTCCACG ATACAGTTGT	360
30	TGCATCTGCT ACCTTGATGC TCGTATTCTT CGGGGCCATT CTTTAAATCT TGGGTCCAGA	420
	CATTATGTCT AATAAAGAAG TCATCACTTC AGGAAGTCTA TTCAATCCTG CTAACAAGA	480
35	TTTCTTTATG TACATTATCC AAACAGCCTT TACCTTCTCA GTTTACTTGT TCGTTTTGAT	540
	GCAAGGTGTC CGAATGTTCT TATCTGAGTT AACAAACGCT TTCCAAGGTA TTTCAAACAA	600
	ATTGTTGCCA GGTTCAATCC CAGCGGTTGA CGTTGCAGCT TCTTATGGAT TTGGTTCTCC	660
40	AAATGCTGTC TTGTCAGGAT TTACCTTTGG TTTGATTGGT CAATTGATTA CAATTGTCTT	720
	GCTCATCGTC TTTAAAAATC CGATTCTTAT TATTACAGGA TTTGTACCAG TGTCTTTGA	780
45	CAATGCAGCC ATTGCGGTCT ACGCTGATAA ACGCGGCGGA TGGAAAGCGG CTGTTATCCT	840
	TTCTTTTATA TCAGGTGTCC TTCAAGTTGC TCTAGGAGCT CTTTGTGTGG CCCTTCTCGA	900
	TTTGGCATCT TATGGTGGCT ACCATGGAAA TATCGACTTT GAATCCCAT GGCTTGGATT	960
50	TGGATATATC TTCAAATACC TTGGTATTGT TGGTTATGTA CTTGTGTGTC TCTTCTTGCT	1020
	TGTTATTCTT CAACTTCAAT TTGCCAAAGC AAAAGATAAA GAGAAATATT ACAACGGTGA	1080
55	AGTTCAAGAA GAAGCTTAGT ATCTAGAAAA GGAGAAATAA AATGGTTAAA GTATTAGCAG	1140
	CGTGCGGAAA TGGAATGGGT TCATCAATGG TTATCAAGAT GAAGGTTGAA AATGCTCTCC	1200
	GTAAGCTTAA TCAAACAGAT TTTACAGTCA ATTCATGCAG TGTCGGTGAA GCTAAAGGTT	1260
60	TAGCAGTAGG ATATGACATC GTAATCGCTT CTCTTCATTT GATTCAAGAA TTGGAAGGGC	1320

- 60 -

	GAACTAATGG GAAGTTAATT GGACTTGATA ACTTGATGGA TGATAAAGAA ATCACCAGAA	1380
	AACTCAGTCA AGCACTACAG TAAAAGGTTG GAGGGGGCTG GACAGAACT GAGAGTTATC	1440
5	GTTTCTGTCC TTCTCCCTCT TTAAATAAAG GAGGCAGATA TGAATTTAAA ACAAGCTTTA	1500
	ATTGACAATG ACTCGATCCG ACTAGGTTTA GAGGCTAACA ATTGGAAAGA AGCAGTCAAG	1560
10	GTAGCAGTAG ATCCCTTAAT TGAAAGTGGG GCAATTTTGC CAGAGTATTA CGATGCTATC	1620
	ATTGAATCGA CTGAAGAGTA TGGGCCTTAC TATATCTTGA TGCCAGGTAT GGCTATGCCC	1680
	CACGCTAGAC CTGAAGCTGG TGTGCAAAGT GATGCCTTTT CATTGATTAC CTTACAAAAT	1740
15	CCTGTTGTAT TTTCAGATGG GAAAGAGGTA TCTGTTTTGT TGGCACTAGC AGCAACAAGC	1800
	TCAAAAATTC ACACAAGTGT AGCCATTCCA CAAATTATG CCCTGTTTGA ATTAGAAGAT	1860
20	TCTATTGCAC GTTTACAGGC TTGCCAGACT AAAGAAGATG TCTTGGCTAT GATTGAAGAA	1920
	TCTAAGGATA GCCCTTATCT CGAAGGATTG GATTTGGAAA GTTAGAAAGA GGAATAAAGA	1980
	AATGACAAAA AGAATACCTA ATTTACAAGT TGCATTAGAC CATTGAGACT TGCAAGGAGC	2040
25	GATTAAAGCA GCTGTTTCTG TTGGTCAGGA AGTAGATATT ATCGAAGCTG GAACTGTTTG	2100
	CTTGCTTCAA GTTGAAGTG AACTGGCTGA AGTCTTGCGT AGCCTTTTCC CAGATAAGAT	2160
30	TATTGTGGCA GACACAAAAT GTGCTGATGC TGGTGGAAAC GTTGCTAAAA ATAATGCGGT	2220
	TCGTGGAGCA GACTGGATGA CTTGTATCTG TTGTGCAACC ATCCCTACTA TGGAAGCAGC	2280
	TCTAAAGGCT ATCAAGACTG AACGAGGAGA ACGAGGCGAA ATCCAGATCG AGCTTTATGG	2340
35	CGATTGGAAT TTTGAACAAG CTCAGCTTTG GCTAGATGCA GGTATTTTAC AAGCTATTTA	2400
	TCACCAATCT CGTGATGCTC TTCTTGCTGG TGAAACTTGG GGTGAAAAAG ACCTTAATAA	2460
40	GGTTAAAAAA CTCATTGACA TGGGCTTCCG TGTATCTGTA ACAGGTGGTC TAGATGTAGA	2520
	TACTCTCAA CTCTTTGAAG GTGTTGATGT CTTTACCTTT ATCGCAGGTC GTGGAATTAC	2580
	AGAGGCTGCG GATCCAGCAG GAGCAGCGCG TGCCTTCAAG GATGAAATCA AACGAATTTG	2640
45	GGGGTAAATC ATGGTACGTC CAATTGGAAT TTATGAAAAG GCAACCCCAA CACACTTTAC	2700
	TTGGCTAGAA CGTTTAAATT TTGCCAAGGA GTTAGGCTTT GATTTTGTCG AGATGTCTAT	2760
50	TGACGAACGT GACGAGCGTT TAGCAAGACT TGAAGGAGT AAGGAAGAAC GCTTGGAAGT	2820
	TGTCAAAGCA ATCTATGAAA CTGGTGTTCG TATTCCTTCT ATCTGTTTTT CAGGCCATCG	2880
	TCGCTACCCA TTGGGTTCAA AAGATCCAGT TCTAGAGGAA AAATCTCTAG AACTCATGAA	2940
55	AAAATGTATC GAATTAGCTC AAGACTTGGG AGTTCGTACG ATTCAATTAG CTGGTTACGA	3000
	TGTTTACTAT GAGGAAAAGT CACCCAGAC ACGCCAACGT TTTATCAAAA ATTTGAGAAA	3060
60	AGCCTGTGAC TGGGCTGAAG AAGCTCAGGT GGTACTTGCT ATTGAAATTA TGGATGATCC	3120
	TTTCATCAAT AGCATCGAAA AATATTTGGC TATAGAAAAA GAGATTGACT CTCCCTTCCT	3180

-61-

CTTTGTATAT CCAGATATTG GTAATGTGTC TGCATGGCAT AATGATATCT ATAGTGAGTT 3240  
 TTATCTTGGT CATCATGCCA TCGCAGCTCT CCATCTCAAG GATACTTATG CAGTGACAGA 3300  
 5 AAGTTCAAAG GGCCAGTTCC GAGATGTACC TTTCCGGCAA GGTGTGTCA AATGGGAAGA 3360  
 AGCTTTCGAT ATTTTAAAGG AAACCAATTA TAATGGACCT TTCCTAATCG AAATGTGGTC 3420  
 10 TGAAAATTGT GAAACAGTAG AAGAAACACG CGCAGCCGTT CAAGAGGCGC AAGCTTTTCT 3480  
 CTATCCACTC ATTAAGAAAG CAGGTTTGAT GTAAGATGAA TCAAGTAATC AATGCTATGC 3540  
 GTAAACGAGT CTGTGATGCC AATCAATCAT TGCCAAAACA TGGACTTGTC AAATTTACCT 3600  
 15 GGGGGAATGT ATCTGAAGTT AATCGCGAAC TCGGTGTCAT TGTATCAAA CCATCAGGCG 3660  
 TGGATTATGA CGAATTGACA CCTGAAAACA TGGTAGTGAC TGATCTAGAT GGTAAGATCC 3720  
 20 CCGGT 3725

## (2) INFORMATION FOR SEQ ID NO:21:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2483 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

35 (iv) ANTI-SENSE: NO

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TCTAGAATCA TTTCCCAGCA GTTGGCTCAG GAAGTCGCAA TTATCTGGGT GAGTTTTTCAG 60  
 CGTGTTGGAC TGAAGTGGAG ATTGATGAGA TATACCGCGC CTTTGTCATG GCACATTTCA 120  
 45 AGAGTTCGCG TCCAGATGCC CAGACCTTGA TTTTCTATAC CCACTATGAC ACTGTGCCAG 180  
 CGGATGGGGA TCAGGTCTGG ACAGAGGATC CTTTACGCT TTCGGTCCGC AATGGCTCAT 240  
 GTATGGGCGT GGGGTTGATG ACGACAGGGT CATATCACAG CTCGCTTGAG TGCTTGAGAA 300  
 50 AATATATGCA GCCCTGATGA TTACCTGTCA ATATCAGCTT TATCATGGAG GGAGCGGAGG 360  
 AATCGGCTTC AACAGACCTA GATAAGTATT TGGAAAAGCA TGCAGACAAA CTCCGTGGGG 420  
 55 CGGATTTGTT GGTCTGGGAA CAAGGGACCA AAAATGCCTT GGAACAGCTG GAAATTTCTG 480  
 GTGGCAATAA GGGGATTGTG ACCTTTGATG CCAAGGTAAA AAGCGCTGAT GTGGATATCC 540  
 ACTCGAGTTA TGGTGGTGTT GTGGAATCAG CTCCTTGGA TCTCCTCCAA GCCTTACAGT 600  
 60 CTCTTCGTGC TGCGGATGGC CGTATCTTGG TTGAAGGCTT GTACGAAGAA GTACAAGAGC 660

- 62 -

	CCAATGAACG AGAAATGGCC TTGCTAGAAA CTTATGGTCA ACGAAACCCA GAGGAAGTTA	720
	GTCGGATTTA TGGATTGGAG TTGCCTCTCT TACAGGAGGA GCGGATGGCC TTTCTAAAAC	780
5	GTTTCTTTTT CGAGCCAGCG CTTAATATCG AAGGAATCCA GTCTGGTTAT CAAGGTCAGG	840
	GTGTTAAGAC TATTTTGCCT GCAGAAGCCA GTGCCAAGCT AGAGGTTCGT CTGGTTCCGG	900
10	GCCTAGAACC GCATGATGTT CTGGAAAAA TTCGGAACA GCTAGACAAA AATGGCTTGG	960
	ATAAGGTAGA ATTATACTAT ACCTTGGGAG AGATACTAGA GTCGAAGCGA TATGAGCGCA	1020
	CCAGCCATTC TCAATGTGAT CGAGTTGGCC AAGAAATTCT ATCCACAGGG CGTTTCAGTC	1080
15	TTGCCGACGA CAGCGGGGAC AGGACCTATG CATACGGTCT TTGATGCCCT AGAGGTACCA	1140
	ATGGTTGCAT TCGGTCTAGG AAATGCCAAT AGCCGAGACC ACGGTGGAGA TGAAAATGTG	1200
20	CGAATCGCTG ATTATTACAC CCATATCGAA TTAGTAGAGG AGCTGATTAG AAGCTATGAG	1260
	TAGAGATATT ATCAAGTTAG ATCAGATCGA TGTGACTTTT CACCAAAGA AGAGAACCAT	1320
	CACAGCGGTT AAGGATGTGA CCATTCACAT CCAAGAAGGG GATATCTACG GAATCGTTGG	1380
25	ATATTCTGGA GCAGGGAAAT CAACCCTTGT ACGGGTGATT AACCTCTGTC AAAAACCATC	1440
	TGCAGGGAAA ATTACCATTG ACGACGATGT GATTTTTGAC GGCAAGGTGA CCTTGACGGC	1500
30	AGAGCAGTTG CGTCGTAAAC GTCAAGATAT CGGGATGATT TTCCAGCATT TTAACCTGAT	1560
	GAGCCAAAAG ACAGCAGAGG AGAATGTAGC CTTTGCCCTT AAACACTCTG GACTCAGCAA	1620
	GGAAGAAAAG AAGGCTAAAG TAGCTAAGTT GTTGACTTGG GTTGGTTTGG CAGATCGTGC	1680
35	TGAAACTAC CCTTCACAAC TATCTGGAGG GCAAAAACAG CGTGTGGCAA TTGCGCGTGC	1740
	CTTGCCAAT GATCCAAAA TCTTGATTTC AGACGAGTCA ACTTCTGCCC TTGACCCTAA	1800
40	GACAACCAAG CAGATTTTGG CCTTGTTGCA AGATTTGAAC CAAAATTAG GATTGACAGT	1860
	TGTCTTGATT ACGCATGAAA TGAGATTGTC AAAGACATTG CCAACCGTGT GCGGTTATG	1920
	CAGGATGGGC ATTTGATTGA AGAGGGAAGT GTCCTTGAAA TCTTCTCAA CCCTAAACAA	1980
45	CCTTTGACTC AAGACTTTAT CTCAACAGCC ACAGGTATTG ACGAAGCCAT GGTCAAATC	2040
	GAGAAGCAAG AAATCGTGGA ACACTTGTCT GAAAACAGTC TCTTGGTGCA ACTTCAAGTA	2100
50	CGCTGGAGCT TCAACAGACG AGCCACTTTT GAATGAATTG TACAAGCATT ACCAAGTAAT	2160
	GGCTAATATT CTCTATGGGA ATATCGAAAT TCTCGATGGT ACTCCTGTTG GAGGAATTGG	2220
	TGGTGGTCTT GTCAGGTGAA AAAGCAGCGT TGGCAGGTGC CCAAGAAGCC ATTCGTCAAG	2280
55	CAGGTGTACA ACTAAAAGTA TTGAAGGGAG TACAGTAAGA TGGAATCATT GATTCAAACC	2340
	TATTTACCAA ATGTCTATAA AATGGGTTGG GCTGTCAGGC AGGCTGGGGG ACGGCTATCT	2400
60	ACTTAACTCT TTATATGCAG TTCTTTCCTT CATTATTCGG GTTCTTGGGG CTAGTGGCAG	2460

-63-

GTCTTCTCGT CTTAAGCGCC AGT

2483

## (2) INFORMATION FOR SEQ ID NO:22:

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1010 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCAATTAATG	TGAGTTAGCT	CACTCATTAG	GCACCCAGG	CTTTACACTT	TATGCTTCCG	60
GCTCGTATGT	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTT	ACACAGGAAA	CAGCTATGAC	120
25 CATGATTACG	CCAAGCTTGC	ATGCCTGCAG	GTCGACTCTA	GAGGATCCAA	GCCATAGTTA	180
GACATGACTG	CCAAATCTAA	GGTTTGAGCA	GTTGTAAAT	AAGCATTAGC	TGTCGCCTCC	240
30 ATGTTGGGAC	TGGTTACTTT	GAGGCCTACT	AAGGCTAGAG	ATCCCAACAT	CATCAGGATC	300
AAGATGGATA	AAAAACGCCC	TTGGAGCCTG	TGAAGGACTG	AATTAAGTCC	TTCCGAATAA	360
GTTTTTCGCT	TGATCATGCT	AGTACTCCAA	ACTGTCAATA	TCCTGAGGAT	GCTGGTTGAG	420
35 CACCACATCC	TTGACACTGG	CATCGTGCAT	TTGAATCACG	CGATCAGCAA	TGGGCGCCAA	480
AGCTCCATTA	TGAGTCACGA	TGATCACCGT	CGCTCCCTTT	TGACGAGACA	TGTCTTGGAG	540
40 AATTTTCAAA	ACCTGCTTGC	CCGTCTGATA	ATCCAAGGCT	CCAGTCGGTT	CATCACAAAG	600
GAGAATTTTA	GGATTTTGG	CTACCGCGCG	TGCAATGGAG	ACTCGCTGTT	GCTCCCCTCC	660
AGAAAGCTGG	GCTGGAAAGT	TATTTAGACG	ATGAGCCAGA	CCTACATCTG	TCAAGACCTG	720
45 ATCAGAATTC	AAGGCATCTG	TCACAATTTT	AGAAGCAGTT	CCACATTTTC	CTTAGCTGTC	780
AGATTAGAAA	CTAGATTATA	AAACTGAAAA	ACAAACCCCA	CATCATTTCT	ACGGTAATTG	840
50 GTGCGCTGGT	GGGAATATA	ATCCGCAATA	TTAACACCAT	CAATCCAGAT	TTCCCCTTCA	900
TCATTGGTAT	CCATTCCCCC	AAGAAGGTTA	AGAACTGTTG	ACTTGCCTGC	ACCTGAAGCA	960
55 CCAAGGATAA	TAACCAAGTTC	CCCCTTTTCA	ATCTCAAAAT	TCACATCACG		1010

## (2) INFORMATION FOR SEQ ID NO:23:

60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1299 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

-64-

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

15	TCGATCGCAC CGTCCCTCTCC TCGTTCTGCT CTTGCTGGGC TATAGTTCCC TCTTCTAGTC	60
	TTGATTTTCT TTGCCCATGC GTTCTTACCA CTTCTACTGT TTGCAGGTTT TACATGTCTG	120
	GATATACTAT TTGTGCTAGG CTTAGCTTCT AGGATGGAGA AAAGAAGTCT AGTAGAGTTA	180
20	TTGAAAGGGG GCATCTTATG ATTGAGTTGA AAAATATTAC CAAAACCATT GGGGGAAAAG	240
	TGATTTTGGG TAACTTATCT CTCAGGATTG ATCAGGGGGA TTTGGTAGCT ATTGTTGGTA	300
	AGAGTGGTAG TGGGAAGTCG ACCTTGTTAA ATTTATTGGG TTTGATAGAT GGTGATTATA	360
25	GCGGACGGTA TGAGATTTT GTTCAGACAA ATCTAGCGGT TAATTCTGCT AAGTCGCAAA	420
	CAATAATCCG TGAACATATC TCTTATCTGT TTCAAATTT TGCCCTGATT GATGATGAAA	480
30	CGGTCGAGTA CAATCTCATG CTGGCGCTGA AATATGTGAA ATTGCCTAAG AAAGACAAGC	540
	TCAAAAAGGT GGAAGAGATT TTAGAGAGAG TAGGTTTGTC AGCTACTTTG CATCAAAGGG	600
	TCTCCGAGTT GTCTGGGGGC GAACAACAAC GAATTGCAGT TGCTAGAGCC ATCTTAAAAC	660
35	CCAGCCAGCT GATTTTAGCC GATGAACCTA CAGGTCGCT GGATCCTGAA AATAGAGATT	720
	TGGTCTTGAA GTTCTCTTA GAGATGAATC GAGAAGGGAA AACAGTCATT ATTGTGACCC	780
40	ACGATGCTTA TGTAGCCCAA CAATGTCATC GTGTCATTGA ATTGGGCGAG GGAAAATGAG	840
	TTCATTGAGC TCCTTTTGAC TGGCTGAATA CTCATGTTTT CCAGAGAAAA ATAGCATAAA	900
	TACGCCTAGG AATGACATT TATGTAGCAT TTCTAGGTTT TTTTGTTC AATTGAAAAT	960
45	TTTTTCAATT TAGGCTTGAC AAAGGATGAG TATAGGAGTA TTATTTATAC AATAAAAAAG	1020
	AATAAACATA AAGAAGGCTT TGTTATGAAT AAGATGAAGA AGGTGTTGAT GACGATGTTT	1080
50	GGTTTAGTGA TGCTCCCCCT ACTATTGCT TGTAAGTAAACA ATCAATCGGC TGGAATTGAA	1140
	GCCATCAAGT CCAAAGGAAA ATTGGTTGTA GCCCTCAATC CAGATTTTGC TCCATTTGAA	1200
55	TATCAAAAAG TGTTTGATGG GAAAAATCAG ATTGTGGGTT CAGATATCGA CTTAGCCAAG	1260
	CTATCGCAAC AGAACTAGGT GTCGACTCTA GAGGATCCC	1299

(2) INFORMATION FOR SEQ ID NO:24:

60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 252 base pairs

-66-

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 913 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AAACACTGCT TCTTGAGCGA ATGACGCTTT GTCCTTTTAA TGAGGTTACC AACGGCTTCA 60  
AAGAGGATTC CCAGCTCGTT CAGCTGTGGA GGTAGCTCGT CTCCTCGTG ATGTAAAAGT 120  
CGAAATTGAA GTCATCGCAG AGATTGGATA AGCTAGTTGA AGTTTGGTGT TGCCAAACTT 180  
CTTTTGATAT AAGGAGAAAA AGATGACAAA GAAACAACTT CACTTGGTGA TTGTGACAGG 240  
GATGGGTGGC GCAGGGAAAA CTGTAGCCAT TCAGTCCTTC GAGGATCTAG GTTATTTTAC 300  
CATTGATAAT ATGCCGCCAG CTCTCTTGCC TAAGTTTTTG CAGCTGGTTG AAATTAAGGA 360  
AGACAATCCT AAGTTGGCCT TGGTAGTGGA TATGCGTAGT CGTTCTTTCT TTTCAGAGAT 420  
TCAAGCTGTT TTGGATGAGT TGGAAAATCA AGATGGTTTG GATTTCAAAA TCCTCTTTTT 480  
GGATGCGGCT GATAAGGAAT TGGTCGCTCG TTACAAGGAA ACCAGACGGA GTCACCCACT 540  
AGCAGCAGAC GGTCTATTT TAGATGGAAT CAAGTTGGAA CGTGAACCTT TGGCACCTTT 600  
GAAAAATATG AGCCAAAATG TGGTGGATAC GACTGAACTC ACTCCACGTG AGCTGCGCAA 660  
AACCCTTGCA GAGCAGTTTT CAGACCAAGA ACAAGCTCAG TCTTTCCGTA TCGAAGTCAT 720  
GTCTTTCGGA TTTAAGTATG GAATCCCGAT TGATGCGGAC TTGGTCTTTG ATGTCCGTTT 780  
CTTGCCAAAT CCCTATTATT TACCAGAACT GAGAAACCAA ACGGGTGTGG ATGAACCTGT 840  
TTATGATTAT GTCATGAACC ATCCTGAGTC AGAAGACTTT TATCAACATT TATTGGCCTT 900  
GATTGAGCCG ATT 913

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5919 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

-67-

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

10	TCGATTCGTG GAGCAGGAAA TCTTTTAGGA AAATCCCAGT CTGGTTTCAT TGATTCTGTT	60
	GGTTTTGAAT TGTATTCGCA GTTATTAGAG GAAGCTATTG CTAAACGAAA CGGTAATGCT	120
	AACGCTAACA CAAGAACCAA AGGGAATGCT GAGTTGATTT TGCAAATTGA TGCCTATCTT	180
15	CCTGATACTT ATATTTCTGA TCAACGACAT AAGATTGAAA TTTACAAGAA AATTCGTCAA	240
	ATTGACAACC GTGTCAATTA TGAAGAGTTA CAAGAGGAGT TGATAGACCG TTTTGGAGAA	300
	TACCCAGATG TAGTAGCCTA TCTTTTAGAG ATTGGTTTGG TCAAATCATA CTTGGACAAG	360
20	GTCTTTGTTC AACGTGTGGA AAGAAAAGAT AATAAAATTA CAATTCAATT TGAAAAAGTC	420
	ACTCAACGAC TGTTTTTAGC TCAAGATTAT TTTAAAGCTT TATCCGTAAC GAACTTAAAA	480
25	GCAGGCATCG CTGAGAATAA GGGATTAATG GAGCTTGAT TTGATGTCCA AAATAAGAAA	540
	GATTATGAAA TTTTAGAAGG TCTGCTGATT TTTGGAGAAA GTTTATTAGA GATAAAAGAG	600
	TCTAAGGAAA AAAATTCCAT TTGATATTTT TCTTCTATAA AATAGATAAA ATGGTACAAT	660
30	AATAAATTGA GGTAAATAAGG ATGAGATTAG ATAAATATTT AAAAGTATCG CGAATTATCA	720
	AGCGTCGTAC AGTCGCAAAG GAAGTAGCAG ATAAAGGTAG AATCAAGGTT AATGGAATCT	780
35	TGGCCAAAAG TTCAACGGAC TTGAAAGTTA ATGACCAAGT GAAATCGCTT GGCAATAAGT	840
	TGCTGCTTGT AAAGGTACTA GAGATGAAAG ATAGTACAAA AAAAGAAGAT GCAGCAGGAA	900
	TGTATGAAAT TATCAGTGAA ACACGGGTAG AAGAAAATGT CTAAAAATAT TGTACAATTG	960
40	AATAATTCTT TTATTCAAAA TGAATACCAA CGTCGTCGCT ACCTGATGAA AGAACGACAA	1020
	AAACGGAATC GTTTTATGGG AGGGGTATTG ATTTTGATTA TGCTATTATT TATCTTGCCA	1080
45	ACTTTTAATT TAGCGCAGAG TTATCAGCAA TTA CTCCAAA GACGTCAGCA ATTAGCAGAC	1140
	TTGCAAACTC AGTATCAAAC TTTGAGTGAT GAAAAGGATA AGGAGACAGC ATTTGCTACC	1200
	AAGTTGAAAG ATGAAGATTA TGCTGCTAAA TATACACGAG CGAAGTACTA TTATTCTAAG	1260
50	TCGAGGGAAA AAGTTTATAC GATTCCTGAC TTGCTTCAAA GGTGATAAAA TGGAAAATTT	1320
	ATTAGACGTA ATAGAGCAAT TTTTGAGTTT GTCAGATGAA AAGCTGGAAG AATTGGCTGA	1380
55	TAAAAATCAA TTATTGCGTT TACAAGAAGA AAAGGAAAGG AAGAATGCGT AAATTCCTAA	1440
	TTATTTTGTT GCTACCAAGT TTTTGGACCA TTTCAAAAGT CGTTAGCACA GAAAAAGAAG	1500
	TCGTCTATAC TTCGAAAGAA ATTTATTACC TTTCACAATC TGACTTTGGT ATTTATTTTA	1560
60	GAGAAAAATT AAGTTCTCCC ATGGTTTATG GAGAGGTTCC TGTTTATGCG AATGAAGATT	1620



-68-

	TAGTAGTGGA ATCTGGGAAA TTGACTCCCA AAACAAGTTT TCAAATAACC GAGTGGCGCT	1680
	TAAATAAACA AGGAATTCCA GTATTTAAGC TATCAAATCA TCAATTTATA GCTGCCGACA	1740
5	AACGATTTTT ATATGATCAA TCAGAGGTAA CTCCAACAAT AAAAAAGTA TGGTTAGAAT	1800
	CTGACTTTAA ACTGTACAAT AGTCCTTATG ATTTAAAAGA AGTGAAATCA TCCTTATCAG	1860
10	CTTATTCGCA AGTATCAATC GACAAGACCA TGTTTGTAGA AGGAAGAGAA TTTCTACATA	1920
	TTGATCAGGC TGGATGGGTA GCTAAAGAAT CAACTTCTGA AGAAGATAAT CGGATGAGTA	1980
	AAGTTCAAGA AATGTTATCT GAAAAATATC AGAAAGATTC TTTCTCTATT TATGTTAAGC	2040
15	AACTGACTAC TGGAAAAGAA GCTGGTATCA ATCAAGATGA AAAGATGTAT GCAGCCAGCG	2100
	TTTGAAACT CTCTTATCTC TATTATACGC AAGAAAAAA TAAATGAGGG TCTTTATCAG	2160
20	TTAGATACGA CTGTAAAATA CGTATCTGCA GTCAATGATT TTCCAGGTTT TTATAAACCA	2220
	GAGGGAAGTG GTAGTCTTCC TAAAAAGAA GATAATAAAG AATATTCTTT AAAGGATTTA	2280
	ATTACGAAAG TATCAAAAGA ATCTGATAAT GTAGTCATA ATCTATTGGG ATATTACATT	2340
25	TCAAACCAAT CTGATGCCAC ATTCAAATCC AAGATGTCTG CCATTATGGG AGATGATTGG	2400
	GATCCAAAAG AAAAATTGAT TTCTTCTAAG ATGGCCGGGA AGTTTATGGA AGCTATTTAT	2460
30	AATCAAAATG GATTTGTGCT AGAGTCTTTG ACTAAACAG ATTTTGATAG TCAGCGAATT	2520
	GCCAAAGGTG TTTCTGTAA AGTAGCTCAT AAAATTGGAG ATGCGGATGG ATTTAAGCAT	2580
	GATACGGGTG TTGTCTATGC AGATCCTCCA TTTATTCTTT CTATTTTCAC TAAGAATTCT	2640
35	GATTATGATA CGATTTCTAA GATAGCCAAG GATGTTTATG AGGTTCTAAA ATGAGGGAAC	2700
	CAGATTTTTT AAATCATTTT CTCAAGAAGG GATATTTCAA AAAGCATGCT AAGGCGGTTT	2760
40	TAGCTCTTTC TGGTGGATTA GATTCCATGT TTCTATTTAA GGTATTGTCT ACTTATCAAA	2820
	AAGAGTTAGA GATTGAATTG ATTCTAGCTC ATGTGAATCA TAAGCAGAGA ATTGAATCAG	2880
	ATTGGGAAGA AAAGGAATTA AGGAAGTTGG CTGCTGAAGC AGAGCTTCCT ATTTATATCA	2940
45	GCAATTTTTT AGGAGAATTT TCAGAAGCGC GTGCACGAAA TTTTCGTTAT GATTTTTTTC	3000
	AAGAGGTCCA TGAAAAAGAC AGGTGCGACA GCTTTAGTCA CTGCCCACCA TGCTGATGAT	3060
50	CAGGTGGAAG CGATTTTTAT GCGCTTGATT CGAGGAACCT CCTTGCCTA TCTATCAGGA	3120
	ATTAAGGAGA AGCAAGTAGT CGGAGAGATA GAAATCATTC GTCCCTTCTT GCATTTTCAG	3180
	AAAAAAGACT TTCCATCAAT TTTTCACTTT GAAGATACAT CAAATCAGGA GAATCATTAT	3240
55	TTTCGAAATC GTATTCGAAA TTCTTACTTA CCAGAATTGG AAAAAGAAAA TCCTCGATTT	3300
	AGGGATGCAA TCCTTAGGCA TTGGCAATGA AATTTTAGAT TATGATTTGG CAATAGCTGA	3360
60	ATTATCTAAC AATATTAATG TGGAAGATTT ACAGCAGTTA TTTTCTTACT CTGAGTCTAC	3420

-69-

	ACAAAGAGTT TTACTTCAAA CTTATCTGAA TCGTTTTCCA GATTTGAATC TTACAAAAGC	3480
	TCAGTTTGCT GAAGTTCAGC AGATTTTAAA ATTTAAAAGC CAGTATCGTC ATCCGATTAA	3540
5	AAATGGCTAT GAATTGATAA AAGAGTACCA ACAGTTTCAG ATTTGTAAAA TCAGTCCGCA	3600
	GGCTGATGAA AAGGAAGATG AACTTGTGTT ACACTATCAA AATCAGGTAG CTTATCAAGG	3660
	ATATTTATTT TCTTTTGGAC TTCCATTAGA AGGTGAATTA ATTCAACAAA TACCTGTTTC	3720
10	ACGTGAAACA TCCATACACA TTCGTCATCG AAAACAGGA GATGTTTTGA TTAAAAATGG	3780
	GCATAGAAAA AAACTCAGAC GTTTATTTAT TGATTTGAAA ATCCCTATGG AAAAGAGAAA	3840
15	CTCTGCTCTT ATTATTGAGC AATTGGTGA AATTGTCTCA ATTTTGGGAA TTGCGACCAA	3900
	TAATTTGAGT AAAAAACGA AAAATGATAT AATGAACACT GTACTTTATA TAGAAAAAT	3960
	AGATAGGTAA AAAATGTTAG AAAACGATAT TAAAAAGTC CTCGTTTCAC ACGATGAAAT	4020
20	TACAGAAGCA GCTAAAAAC TAGTGCTCA ATTAACATA GACTATGCAG GAAAAATCC	4080
	AATCTTAGTT GGGATTTTAA AAGGATCTAT TCCTTTTATG GCTGAATTGG TCAAACATAT	4140
25	TGATACACAT ATTGAAATGG ACTTCATGAT GGTTCCTAGC TACCATGGTG GAACAGCAAG	4200
	TAGTGGTGTT ATCAATATTA AACAAGATGT GACTCAAGAT ATCAAAGGAA GACATGTTCT	4260
	ATTTGTAGAA GATATCATTG ATACAGGTCA AACTTTGAAG AATTTGCGAG ATATGTTTAA	4320
30	AGAAAGAGAA GCAGCTTCTG TTAAAATTGC AACCTTGTTG GATAAACCAG AAGGACGTGT	4380
	TGTAGAAATT GAGGCAGACT ATACCTGCTT TACTATCCCA AATGAGTTTG TAGTAGGTTA	4440
35	TGGTTTAGAC TACAAAGAAA ATTATCGTAA TCTTCCTTAT ATTGGAGTAT TGAAAGAGGA	4500
	AGTGTATTCA AATTAGAAAG AATAATCTTT AATGAAAAA CAAAATAATG GTTTAATTAA	4560
	AAATCCTTTT CTATGGTTAT TATTTATCTT TTTCTTGTG ACAGGATTCC AGTATTTCTT	4620
40	ATTCTGGGAA TAACTCAGGA GGAAGTCAGC AAATCAACTA TACTGAGTTG GTACAAGAAA	4680
	TTACCGATGG TAATGTAAAA GAATTAACCT ACCAACCAAA TGGTAGTGTT TCGAAGTTTC	4740
45	TGGTGTCTAT AAAAATCCTA AAACAAGTAA AGAAGGAACA GGTATTCAGT TTTTCACGCC	4800
	ATCTGTTACT AAGGTAGAGA AATTTACCAG CACTATTCTT CCTGCAGATA CTACCGTATC	4860
	AGAATTGCA. AAACCTGCTA CTGACCATAA AGCAGAAGTA ACTGTTAAGC ATGAAAGTTC	4920
50	AAGTGGTATA TGGATTAATC TACTCGTATC CATTGTGCCA TTTGGAATTC TATTCTTCTT	4980
	CCTATTCTCT ATGATGGGAA ATATGGGAGG AGGCAATGGC CGTAATCCAA TGAGTTTTGG	5040
55	ACGTAGTAAG GCTAAAGCAG CAAATAAAGA AGATATTAAA GTAAGATTTT CAGATGTTGC	5100
	TGGAGCTGAG GAAGAAAAAC AAGAACTAGT TGAAGTTGTT GAGTTCTTAA AAGATCCAAA	5160
	ACGATTCACA AAACCTGGAG CCCGTATTCC AGCAGGTGTT CTTTGGAGG GACCTCCGGG	5220
60	GACAGGTAAG ACTTTGCTTG CTAAGGCAGT CGCTGGAGAA GCAGGTGTTT CATTCTTTAG	5280

-70-

5 TATCTCAGGT TCTGACTTTG TAGAAATGTT TGTCGGAGTT GGAGCTAGTC GTGTTTCGCTC 5340  
 TCTTTTGTGAG GATGCCAAAA AAGCAGCACC AGCTATCATC TTTATCGACT GAAATGGATG 5400  
 CCCGTGGGAC GTCAACGTGG AGTCGGTCTC GCGGAGGTA ATGACGAACG TGAACAAACC 5460  
 TTGAACCAAC TTTTGATTGA GATGGATGGT TTTGAGGGAA ATGAAGGGAT TATCGTCATC 5520  
 10 GCTGCGACAA ACCGTTTCTG TGTAAGTGGT CCTGCCCTTT TGCCTCCAGG ACCTTTTGGT 5580  
 AGAAAAGTAT TGGTTGGCCG TCCTGATGTT AAAGGTCGTG AAGCAATCTT GAAAGTTCAC 5640  
 15 GCTAAGAAC AAGCCTTTAGC AGAAGATGTT GATTTGAAAT TAGTGGCTCA ACAAACCTCA 5700  
 GGCTTTGTG GTGCTGATTT AGAGAATGTC TTGAATGAAG CAGCTTTAGT TGCTGCTCGT 5760  
 CGCAATAAAT CGATAATTGA TGCTTCAGAT ATGATGAAAG CAGAAGATAG AGTTATTGCT 5820  
 20 GGACCTTCTA AGAAAGATAA GACAGTTTCA CAAAAGAAC GAGAATTGGT TGCTTACCAT 5880  
 GAGGCAGGAC ATACCATTGT TGGTCTAGTC TTGTCGACT 5919

## (2) INFORMATION FOR SEQ ID NO:28:

- 25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1863 base pairs  
 (B) TYPE: nucleic acid  
 30 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (iii) HYPOTHETICAL: NO  
 35 (iv) ANTI-SENSE: NO

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GAGCTCGGTA CCCGGGGATC ATACTCAAGA GGAGGTAATC CAATGAACAC TAGTCTTAAA 60  
 45 CTCAGCAAAC AACTCAGTTT TGGAGAGGAG ATTGCTAATA GCGTGACCCA TGCTGTGGGT 120  
 GCAGTCATCA TGCTTATCTT GCTGCCTATT TCATCCATCT ATAGTTATGA AGCACACGGA 180  
 50 TTTTATCCT CTATCGGCGT TTCCATTTTC GTCATCAGTC TCTTTCTCAT GTTCCTATCA 240  
 TCCACCATT ATCACTCTAT GGCCTATGGT TCGACCCACA AATATGTTTT GCGAATCATT 300  
 GACCATTCTA TGATTTACGT TGCCATTGCC GGCTCATACA CGCCCGTTGT CTTGACCTTG 360  
 55 ATGAATAACT GGTGTTGGTA TCTGATTATT GTCATCCAAT GGGGAACGAC CATCTTTGGT 420  
 ATTCTCTATA AAATCTTTGC TAAAAAGGTC AATGAGAAAT TTAGCCTTGC TCTTTACCTG 480  
 ATTATGGGCT GGTGTTTCT GGCTATCATT CCTGCCATTA TCAGTCAAAC GACACCCGTT 540  
 60 TTCTGGAGTC TCATGGTAAC TGGCGGACTC TGTTATACAG TTGGAGCTGG ATTTTATGCC 600

-71-

AAGAAAAAAC CTTATTTCCA CATGATTTGG CATCTCTTTA TCCTAGCTGC GTCCGCACTC 660  
CAATACATCG CTATTGTTTA TTACATGTAA AAAAGTTGAG AAATTCAATC TCAACTTTTT 720  
5 TCTTTACACA TATTGATAAA GTACTGGTGC AAGCGCACAT CATCAGTCAA TTCTGGATGA 780  
AAAGAACTTA CCAACATATT TTTTCTTGG GCTGCAACAA TTTGATTGTT CACTATTGCT 840  
10 AAAATTTCTA CACCCTCACC AACACTACTG ATAATCGGAC CACGGATAAA GGTCATTGGA 900  
ATCTTGCCAA CTCCCTTACA TTCTGCTTCC GTGTAGAAAC TTCCTAATTG GCGCCCATAA 960  
GCATTACGCT CGACCACCAT ATCCATAGTT CCTAGATGAC TCTCTTTCTG AGAAGTGATT 1020  
15 TCCTTAGCCA GCAAATTA GCGCGCACAG GTCCCAAACA CTGGTAAGCC AGATAGAATG 1080  
GCTTCTCGTA TGGGAAGTAG CATGTTCTGG TCACGTAAGA GCTTGCCCAT GGTGTAGAC 1140  
20 TCACCACCAG GCAAATAAA CCCGACAAGT CACTCTGATC TTGCTGAAA CATCTAGATT 1200  
TCTGAGTTCT ACACTCTCGA CACCTAATTG ATCTAGCACT TTTGCATGTT CTGCAAAGGC 1260  
CCCTTGCAAG GCCAATATTC CGATTTTCAT CTATTTTCCT CGTTCAGCCA TGAGAATTTG 1320  
25 GATTCATTTT CATTAAATACC AACCATGGCT TCTCCTAAAT CTTCAGAGAT TTGAGCTAGG 1380  
ATTTGAGGAT TACGGAAGTT AGTCACAGCC TTAACAATGG CACTCGCTCG TTTAACAGGA 1440  
30 TCTCCTGACT TGAAATACC TGAACCGACA AAGACCCCT CTGCCCTAA TTGCATCATT 1500  
AACGCAGCAT CTGCTGGCGT TGCAACACCT CCAGCAGCGA AATTTACAAC TGGCAATTTT 1560  
CCATGTTTCAT GAACATATTG GACCAATTCT ACAGGGACTT GCAAATCCTT GGCAGCAACA 1620  
35 TAAAGCTCGT CCTCACGTAA GTTTTGAATG CGGCGAATTT CCTGATTCAT CATACGCATA 1680  
TGACGAACAG CTTGGACTAT ATCCCCTGTC CCTGGTTCTC CTTTAGTACG AATCATGGAA 1740  
40 GCACCTTCAG CGATACGACG CAAGGCTTCA CCCAAATCCT TAGCACCACA GACAAAAGGA 1800  
ACTTGGAATT CTTTCTTGTC CACATGGAAA CGGTCATCAG CTGGAGATAG AACTTCACTC 1860  
TCG 1863  
45

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5448 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

60

-72-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

5	TAAAGAAGGT GGATTTGAAG TTAACGGTAA ATTCATCAAA GTTTCTGCTG AACGTGATCC	60
	AGAACAAATC GACTGGGCTA CTGACGGTGT AGAAATCGTT CTTGAAGCTA CTGGTTTCTT	120
	TGCTAAGAAA GAAGCAGCTG AAAAACACCT TAAAGGTGGA GCTAAAAAG TTGTTATCAC	180
10	TGCTCCTGGT GGAAACGACG TTA AACAGT TGTATTCAAC ACTAACCACG ACGTTCTTGA	240
	CGGTACTGAA ACAGTTATCT CAGGTGCTTC ATGTACTACA AACTGCTTGG CTCCAATGGC	300
	TAAAGCTCTT CAAGACA ACT TTGGTGTTGT TGAAGGATTG ATGACTACTA TCCACGCTTA	360
15	CACTGGTGAC CAAATGATCC TTGACGGACC ACACCGTGTG GTGACCTTCG CCGTGCTCGC	420
	GCTGGTGCTG CAAACATCGT TCCTAACTCA ACTGGTGCTG CAAAAGCTAT CGGTCTTGTA	480
20	ATCCCAGAAT TGAATGGTAA ACTTGATGGA TCTGCACAAC GCGTTCCAAC TCCA ACTGGA	540
	TCAGTTACTG AATTGGTCGC AGTTCTTGAA AAGAACGTTA CTGTTGATGA AGTGAACGCA	600
	GCTATGAAAG CAGCTTCAAA CGAATCATA GGTACACAG AAGATCCAAT CGTATCTTCA	660
25	GATATCGTAG GTATGTCTTA CGGTTTATTG TTTGACGCAA CTCAACTAA AGTTCTTGAC	720
	GTTGACGGTA AACAATTGGT TAAAGTTGTA TCATGGTACG ACAACGAAAT GTCATACACT	780
30	GCACA ACTTG TTCGTACTCT TGAATACTTC GCAAAGATTG CTAAATAATT CTTGAGTTGA	840
	TAGAAAGCAA GGCTTTGTGG TCTTGCTTTT TTATATGGAA AAATGGATGA CACGATCATC	900
	CATTCTTTTT TAATTCTTTT TCAAATGTAT TTGAAAGGGT AGTGAAAGTT AGCCTCTCTA	960
35	AAGTAAGTGG GTGGGTAAAG GAAAGTCGGA AGGCATGAAG CATAAGCCGG CTTGTCTTTG	1020
	ATTTACTATT ATAGAGAGGG TCTCCCAGGA TAGGAAGATT ATGATGCAAA AGGTGCACAC	1080
40	GAATCTGATG GGTTGCCCCT GTCTTTAGCT TGCAATGAGC CAAGGAAGTC TTGTTTGAGA	1140
	ATTGCTTTAA TCTGCTTACA TCGGTTTCAG CATATTCCC ATTTTGTGCA TCAACTATTC	1200
	TTTTTCTACG ATCATGGCGA TCACGTCCAA TTTTGTCTCT GAAAACAAGT TCTTTTCTGT	1260
45	TGATATTTCC ATCAACTAGA GCCCAATATT CTCTAGAAAT CTCTTTTTC TCCAATAAGC	1320
	GATTGAGAAT GGGCAGGATA AAAGGATTTT TGGCAAAGAG AACTAAGCCA CTGGTTTCCA	1380
50	TGTCCAGACG ATGAACGACA TAGCAGGTTT GGCCAACATA GGTACTGACA TGGTTAAGAA	1440
	GGGCAATTTT GTTTGGTTGA TTACCATGCG TTTTCATCCC CTCTGGTTTG TTTACAATAA	1500
	TCAAGTGTTG ATCTTGATAA ACTTCCTGCA CTAAGTCTGG GTTGCCCCAA GGGATCGTCT	1560
55	TTTGGGGATA ATCTTCCTCG TCAAAAGTCA ACTGGCAAAC ATCTCCAGGA TTTACGATTT	1620
	CGTTCCAGCG GACTTCTTCT TGATTTATCA AAATATGTTT CTTGATTCTC AAAAAATGAC	1680
60	GGATTTTCT AGGGATGAGG AGTTGTTTCT CAAGTAATTG CTTTACCGTC ATTTGAGGTA	1740

-73-

	GAGAGGCGGG	TAATGTAAAT	GTGAATTGCA	TACAGATATT	GTAACAAAAA	AAGCCCTATT	1800
	TGGATAGGAA	ATAGCTAAAT	TCTTGTCTTC	CTATGATGAA	GATGATAAAA	TAAACGCATG	1860
5	AAATTAGATA	AATTATTGA	GAAATTTCTT	TCTCTTTTAA	AAAAAGAAAC	AAGTGAACATA	1920
	GAGGACTCTG	ATTCTACTAT	CTTACGTCGC	TCTCGTAGTG	ATCGAAAAAA	ATTAGCCCAA	1980
	GTAGGTCCGA	TTCGAAAATT	CTGGCGTCGT	TATCATCTAA	CAAAGATTAT	CCTTATACTA	2040
10	GGTTTGAGTG	CAGGCTTGCT	AGTTGGAATC	TATTTGTTTG	CTGTAGCCAA	GTCGACCAAT	2100
	GTCAATGATT	TGCAAAATGC	CTTGAAAAC	CGGACTCTTA	TTTTTGACCG	TGAAGAAAAA	2160
15	GAGGCTGGTG	CCTTGTCTGG	TCAAAAGGGA	ACCTATGTTG	AGCTGACTGA	CATCAGTAAA	2220
	AACTTGACGA	ATGCTGTTAT	TGCGACAGAA	GACCGTTCTT	TCTATAAAAA	TGACGGGATT	2280
	AACTATGGCC	GTTTCTTCTT	GGCTATTGTC	ACTGCTGGAC	GTTGAGGTGG	TGGCTCTACC	2340
20	ATTACCCAAC	AGCTGGCTAA	AAACGCCTAT	TTATCGCAGG	ATCAAACGTG	TGAGAGAAAA	2400
	GCGAAAGAAT	TTTTCTTCTG	CTTAGAATTA	AGCAAAAAAT	ATAGTAAGGA	GCAAATTCTA	2460
25	ACCATGTACC	TTAACAACGC	TTATTTTGGA	AATGGTGTGT	GGGGTGTAGA	AGATGCGAGT	2520
	AAGAAATACT	TTGGAGTTTC	TGCATCAGAA	GTGAGTCTGG	ATCAAGCTGC	GAATCTGGCA	2580
	GGGATGCTCA	AGGGGCCGGA	ACTGTATAAT	CCCTTGAATT	CCGTAGAAGA	TTCTACTAAT	2640
30	CGGCGCGATA	CTGTCTTGCA	GAATATGGTT	GCAGCAGGAT	ATATTGATAA	AAACCAAGAA	2700
	ACCGAAGCTG	CTGAAGTTGA	TATGACTTCG	CAATTGCACG	ATAAGTATGA	AGGAAAAATC	2760
35	TCAGATTACC	GTTACCCCTC	TTATTTTGAT	GCGGTGGTTA	ATGAAGCTGT	TTCCAAGTAT	2820
	AATCTAACAG	AGGAAGAGAT	TGTCAATAAT	GGCTACCGCA	TTTACACAGA	GCTGGACCAA	2880
	AACTACCAAG	CAAATATGCA	GATTGTTTAT	GAAAACACAT	CGCTATTTC	GAGGGCAGAG	2940
40	GATGGAACGT	TTGCTCAATC	AGGAAGTGTA	GCTCTCGAAC	CGAAAACAGG	GGGAGTTCGT	3000
	GGAGTTGTCG	GTCAAGTTGC	TGACAATGAT	AAAACGGAT	TCCGGAATTT	CAACTATGCA	3060
45	ACCCAATCAA	AGCGTAGTCC	TGGTCTTACA	ATTAAGCCTT	TAGTTGTTTA	TACACCAGCA	3120
	GTTGAAGCAG	GCTGGGCTTT	GAATAAGCAG	TTGGATAACC	ATACCATGCA	GTATGATAGC	3180
	TATAAGGTTG	ATAACTATGC	AGGGATCAAA	ACAACTCGAC	AAGTTCCTAT	GTATCAATCC	3240
50	TTGGCAGAAT	CGCTTAATCT	ACCTGCTGTT	GCCACTGTTA	ATGATTGTTG	TGTTGACAAG	3300
	GCTTTTGAGG	CAGGCGAAAA	ATTCGGACTC	AACATGGAAA	AGGTCGACCG	TGTTCTTGGT	3360
55	GTCGCCTTGG	GAAGCGGTGT	TGAAACCAAC	CCTCTTCAAA	TGGCTCAAGC	ATACGCTGCC	3420
	TTTGCAAATG	AAGGTTTAAT	GCCTGAAGCT	CATTTTATTA	GTAAGATTGA	AAATGCTAGT	3480
	GGACAAGTTA	TTGCGAGTCA	TAAAAATTCA	CAAAAACGGG	TGATTGATAA	GTCTGTAGCT	3540
60	GACAAGATGA	CCAGTATGAT	GTTGGGGACT	TTCACCAACG	GTACCGGTAT	TAGTTCATCG	3600

-74-

	CCTGCAGACT ATGTCATGGC AGGGAAAAC T GGAACAACTG AAGCAGTTTT CAATCCGGAG	3660
	TACACAAGTG ACCAGTGGGT AATTGGTTAT ACTCCGGATG TAGTGATTAG CCACTGGCTT	3720
5	GGCTTTCCGA CCACTGATGA AAATCACTAT CTAGCTGGCT CTACTTCAAA CGGTGCAGCT	3780
	CATGTCTTTA GAAACATTGC CAATACTATT TTACCTTATA CGCCAGGAAG TACCTTTACG	3840
10	GTTGAAAATG CTTATAAGCA AAATGGAATT GCACCAGCCA ATACAAAAAG ACAAGTACAA	3900
	ACCAATGATA ATAGCCAGAC AGATGATAAT TTGTCTGATA TTCGAGGGCG TGCGCAAAGT	3960
	CTAGTAGATG AGGCTAGCCG GGCTATCTCA GATGCGAAGA TTAAGGAAAA GGCTCAAACA	4020
15	ATATGGGATT CGATAGTCAA TCTATTTTCGC TAAGATGCTT GTCAAAGCCT AGCTTTCTTG	4080
	TTATAATGGA TAAGATGGAG GCGTTATGGC ACTAAAAAAA GCAAGCCTAG CTTGTGCGGT	4140
20	TTGTGGTTCG AGAACTATT CAATCAAGAT CAGCGGAAAC CCCAAGCCTA CACGACTAGA	4200
	AGTAAATAAA TTTTGTAAGC ATTGTGGCAA GTACACTACA CACAGAGAAA CGAGATAGGA	4260
	GAGAGCGATG CGTTTTATTG GAGATATTTT TAGACTTCTT AAAGACACAA CATGGCCAAC	4320
25	TCGCAAGGAA AGCTGGAGAG ATTTTCGTTT TATCATGGAA TACACAGCTT TCTTTGTAGT	4380
	AATTATTTAC ATTTTGTACC AGTTGATTGT TTCAGGTTTG ATTCGATTTA TTAACATTTT	4440
30	TTAGAAGATT AGTGGAGTTA ATTACACTAG AAATCTTCTA TTTATGAAAG GAAATATCAT	4500
	GGATAGTTTT GATAAAGGAT GGTTTGTTTT ACAAACTTAT TCTGGTTATG AAAATAAGGT	4560
	GAAAGAAAAT CTATTACAAC GTGCACAAAC CTACAATATG TTGGATAATA TTCTACGCGT	4620
35	TGAAATTCCA ACACAAACAG TGCAAGTTGA AAAAAATGGA AAGAGAAAAG AAGTAGAAGA	4680
	AAATCGCTTT CCAGGTTATG TTCTTGTAGA AATGGTCATG ACAGATGAAG CTTGGTTTGT	4740
40	TGTTCGAAAC GCACAGAGTC CTACAAAATT CATTTTCAGAA CAAACAGCTT ATGAAATTGA	4800
	TGAAGAGGTT CGTTCATTAT TAAATGAGGC ACGAAATAAA GCTGCTGAAA TTATTCAGTC	4860
	AAATCGTGAA ACTCACAGT TAATTGCAGA AGCATTATTG AAATACGAAA CATTGGATAG	4920
45	TACACAAATT AAAGCTCTTT ACGAAACAGG AAAGATGCCT GAAAGCAGTA GAAGAGGAAT	4980
	CTCATGCACT ATCCTATGAT GAAGTAAAGT CAAAAATGAA TGACGAAAAA TAACCCTGAG	5040
50	AGAGGCTGGA GCCTCTCTTT TTTGTGCAGT TTAGGAGCTA AAGGGAACAG AATGGAGAAA	5100
	ATGGAACAAA TGTGTTTTCT AATCTGTTAG ACTGTATCTA GAAAGGGGAA AATTATGATT	5160
	AAAGAATTGT ATGAAGAAGT CCAAGGGACT GTGTATAAGT GTAGAAATGA ATATTACCTT	5220
55	CATTTATGGG AATTGTCGGA TTGGGACCAA GAAGGCATGC TCTGCTTACA TGAATTGATT	5280
	AGTAGAGAAG AAGGACTGGT AGACGATATT CCACGTTTAA GGAAATATTT CAAAACCAAG	5340
60	TTTCGAAATC GAATTTTAGA CTATATCCGT AAGCAGGAAA GTCAGAAGCG TAGATACGAT	5400

-75-

AAAGAACCCT ATGAAGAAGT GGGTGAGATC CCCGGTACCG AGCTCGAA

5448

## (2) INFORMATION FOR SEQ ID NO:30:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1040 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- 10 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 15 (iv) ANTI-SENSE: NO

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TAGTGAATTC	GAGCTCGGTA	CCCGGGGATC	GTTTCTCGGT	TCTTTTGGAG	CACAAGGGCA	60
TCCATCCCAT	TGTCTATATT	TCCAAATGG	ATTGTGTGA	AGATAGGGGA	GAAGTGGATT	120
25 TTTACCAGCA	GACCTATGGT	GACATCGGCT	ATGACTTTGT	GACCAGTAAA	GAGGAACTCC	180
TGTCTTTGTT	AACAGGCAAG	GTTACGGTCT	TTATGGGGCA	GACAGGTGTT	GGGAAGTCAA	240
30 CTCTTCTCAA	TAAAATCGCA	CCAGACCTCA	ATCTTGAAAC	GGGAGAAATT	TCAGACAGTC	300
TAGGTCGCGG	TCGCCATACC	ACTCGAGCTG	TTAGTTTTTA	CAATCTCAAC	GGGGGTAAAA	360
TCGCAGATAC	ACCAGGATTT	TCATCCTTGG	ACTATGAAGT	ATCAAGGGCT	GAAGACCTCA	420
35 ATCAGGCTTT	CCCAGAGATT	GCTACTGTTA	GCCGAGATTG	TAAGTTCCGT	ACTTGTACCC	480
ATACCCATGA	GCCGTCTTGT	GCCGTCAAAC	CAGCTGTTGA	AGAGGGTGTT	ATTGCAACCT	540
40 TCCGTTTTGA	CAATTACCTG	CAATTCCTTA	GTGAAATTGA	AAATCGTAGA	GAAACCTATA	600
AAAAAGTCAG	CAAAAAAATT	CCAAAATAAG	GAGAAACCTA	TGTCTCAATA	CAAGATTGCT	660
CCGTCAATTC	TGGCAGCAGA	TTATGCCAAC	TTTGAACGTG	AAATCAAACG	TCTAGAAGCA	720
45 ACTGGGGCAG	AATATGCCCA	TATCGATTCT	GGACAGTCAT	TTGTACCGC	AAATCAGTTT	780
TGGTGCAGGT	GTGGTCGAGA	GCTTCGTCTT	CATAGTAAGA	TGGTTTTCGA	TTGCCACTTG	840
50 ATGGTGTCAA	ACCCTGAGCA	TCATCTGGAA	GATTTTGCGC	GTGCAGGTGC	AGACATCATC	900
AGTATCCATG	TAGAAGCAAC	ACCTCATATT	CATGGCGCCC	TCCAAAAAAT	TCGTTCATCTC	960
GGAGTTAAGC	CTTCAGTCGT	TATCAATCCT	GGCACACCAG	TTGAAGCCAT	CAAGCACGTC	1020
55 CTTCATCTAG	TGACAAGTTT					1040

## (2) INFORMATION FOR SEQ ID NO:31:

- 60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 789 base pairs



-76-

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

10 (iv) ANTI-SENSE: NO

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ATATCACGAC	GGAGCCATAC	TACCGATTTT	CTTAAGCATA	GCGCCACCTT	TACCGATGAT	60
AATCCCTTTT	TGGCTATCGC	GCTCGACCAT	GATGGTTGCA	CGGATGTGAA	CCTTGTCTGT	120
20 CTCTTCGTCT	CGTTTCATAG	AGTCAACAAC	TACTGCTACA	GAATGCGGAA	TCTCTTCACG	180
AGTTAGGTGC	AAGACTTTCT	CGCGAACCAT	TTCTGAAACT	AAGAAACGTT	CTGGATGATC	240
TGTGATTTGA	TCAGACGGGA	AATATTGGAA	ACCTTCATCC	AGATTTTCAC	TCAAAATATC	300
25 CACTAGACGA	GACACGTTAT	TTCCCTGAAG	GGCTGAGATT	GGAACAATTT	CCTTAAAGTC	360
CATTTGATTA	CGGAAGTCAT	CAATCTGAGA	CAAGAGCTGG	TCTGGATGGA	CCTTATCGAT	420
30 TTTATTACAC	ACCAAATCA	CAGGAACCTT	GGCAGCCTTG	GAGACGCTCG	ATAATCATAT	480
CGTCCCCCTT	ACCACGCGCT	TCATCAGCAG	GCACCATGAA	AAGAACAGTG	TCCACTTCGC	540
GAAGGTACTG	TAGGCAGACT	CAACCATGAA	ATCTCCGAGA	GCTGTTTTAA	GTTTGTGAAT	600
35 CCCTGGTGTG	TCGATAAAGA	CAATTTGCTC	CTTATCAGTC	GTGTAATTCC	CATGATTTTA	660
TTGCGCGTTG	TCTGCGCCTT	GTCACCATG	ATGGCAATCT	TTGCCCCAT	AACGTGATTT	720
40 AAAAAGGTTG	ACTTCCCAAC	ATTGGGACTC	CTAAATGGC	TACAAACCTG	ATTTAAAATT	780
CATAATTCC						789

45 (2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1233 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

55 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

-77-

5 TGAGATAAAC TTGCGACTCA TATGAGAATA TGAATCAAGC CGTCCTCGTG AACCCCGATA 60  
TCAACGAAGG CACCGAAGTC AACAAACGTTA CGCACAAACAC CTTCTAGCTT CTGACCTACC 120  
ACTAGGTCTT TGATATCTAG GACATCTTGG CGAAGCACAG GTGCGTCAAA GGAATCACGG 180  
AAATCTCGAC CTGGTTTGAG AAGATCTGCA ATGATATCTT TAAGGGTTTC TGGACCGAGG 240  
10 TCTAGCTCTT GAGCCATTTC CTTGACTGAA AGGGACTTGA GTTTGCTTGG GCTTCTTCGT 300  
TTAGGTCTTT AATATCTAAA CGTTTGAAGA GTCCTTAACT GCAGTGAAT TCTCTGGGTG 360  
AACTCCTGTA TTATCAAGGA TATTGCTACT TTCAGGGATA CGAAGGAAAC CAGCAGCCTG 420  
15 CTCAAAGGCC TTGGCTCCCA GACGAGGAAC TTTCTTGATT TGGGCGCGTG AAGTGATTTT 480  
TCCTTCTTCC TCGCGGTATT TGACAATATT TTCAGAGATA GTTTTGTTGA GTCCAGCTAC 540  
20 GTGTGAAAGA AGAGCTGGGC TAGCTGTATT GACATTGACA CCAACTTGGT TAACCACTGT 600  
ATCGACAACA AAGTCCAGAC TCTCAGATAG TTTCTTCTGA CTGACATCGT GTTGGTATTG 660  
ACCGACACCA ATTGACTTAG GATCGATTTT GACCAATTCC GCAAGAGGAT CTGCAAACG 720  
25 ACGGGCGATA GAAATGGCAG AGCGTTTTTC AACGGTCAAG TCTGGAAACT CCTGACGAGC 780  
AAGTTCGCTG GCAGAATAGA CAGAAGCACC ACTTTCATTA ACGATAACAT AGCTGACTTC 840  
30 AGGGAAATCT TTCAGAACTT CCGCTACAAA AGCTTCACTT TCACGACTGG CCGTTCCATT 900  
TCCAATGGCA ATAATCTCTA CACCGTATTG ACCAATTAAA TCTGCTAAAT CTTTCTTGGC 960  
TTCTTCGATT TGACGAGCTG ATGCTGGTTT AACAGGATAA ATACCTGAG TTGTCAGCAT 1020  
35 TTTTCTGTG GCATCCACGA CAGCTAACTT GGCACCTGTA CGAAAGGCTG GGTCAAATCC 1080  
AAGAACCACG CGCCCTTTCA GTGGAGCAAC CAAGAGGAGA TTGCGCAGAT TGTCAGAAAA 1140  
40 AAGTTGGATA GCTCCCTCTT CAGCTTTCTC AGTTAATTCT GTCCGAATAC GACGCTCAAT 1200  
AGCAGGCAAG ACCTTTTTTCT TAACGGATTG CTG 1233

## (2) INFORMATION FOR SEQ ID NO:33:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6679 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
55 (iv) ANTI-SENSE: NO

- 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

-78-

	ACAAGGCGTC ATCCGTGTAT TTCTTAAATA CGGACCAAAT GGTGAGAAAG TTATCACTAA	60
5	CTTGAAACGT GTTCTAAAC CAGGACTTCG TGTCTACAAA AAACGTGAAG ACCTTCCAAA	120
	AGTTCTTAAC GGACTTGGA TTGCCATCCT TTCAACTTCT GAAGGTTTGC TTACTGATAA	180
	AGAAGCACGC CAAAAGAATG TTGGTGGTGA GGTATCGCT TACGTTTGGT AAAATCAAGA	240
10	TACAAAGCTC GTAAAGAACA AAGCAAAATT AGGAAGTTGG AGAAGTTTGT TTACAAACAA	300
	GCCAACTTAT CTATTTTGCA CAGTTCTTAG ATCGTGTTCA GTTCAGCTCT TGAATAAAT	360
15	AAGTATCTGA ACCCGTGAA AACTGGCCGT TCTGGCTGAC AATTTAACAG GAGAAAATAA	420
	ACATGTCACG TATTGGTAAT AAAGTTATCG TGTTCCTGC TGGTGTGAA CTCGCTAACA	480
	ATGACAACGT TGTAACGTGA AAAGGACCTA AAGGAGAACT TACTCGTGAG TTCTCAAAAG	540
20	ATATTGAAAT CCGTGTGGAA GGTACTGAAG TAACTCTTCA CCGTCCAAAC GATTCAAAAG	600
	AAATGAAAAC TATCCACGGA ACTACTCGTG CCCTTTTGAA CAACATGGTT GTTGGTGTAT	660
25	CAGAAGGATT CAAGAAAGAA CTTGAAATGC GTGGGGTTGG TTACCGTGCA CAGCTTCAAG	720
	GATCTAAACT TGTTTTGGCT GTTGGTAAAT CTCATCCAGA CGAAGTTGAA GCTCCAGAAG	780
	GAATTACTTT TGAACCTCCA AACCCAACAA CAATCGTTGT TAGCGGAATT TCAAAGAAG	840
30	TAGTTGGTCA AACAGCTGCT TACGTACGTA GCCTTCGTTT ACCAGAACCA TATAAAGGTA	900
	AAGGTATCCG TTACGTTGGT GAATTCGTTT GTCGTAAAGA AGGTAAAACA GGTAAATAAT	960
35	GTTGAGTGGT TGATCATCAA CCACCAACCT ATTTTCCAAC TTTGTGCATA GCAACGATTT	1020
	AAACTAAAG AGGTGAAAAC TGTGATTTC AACCAGATA AAAACAACT CCGCCAAAAA	1080
	CGCCACCGTC GCGTTCGGGA AACTCTCTG GAACTGCTGA TCGCCACGCT TTGAACGTAT	1140
40	TCCGTTCTAA TACAGGCATC TACGCTCAAG TGATTGATGA CGTAGCGGGT GTAACGCTCG	1200
	CAAGTGCTTC AACTCTTGAT AAAGAAGTTT CAAAAGGAAC TAAACTGAA CAAGCCGTTG	1260
45	CTGTCGGTAA ACTCGTTGCA GAACGTGCAA ACGCTAAAGG TATTTAGAA GTGGTGTTCG	1320
	ACCGCGGTGG ATATCTATAT CACGGACGTG TGAAAGCTTT GGCTGATGCA GCTCGTGAAA	1380
	ACGGATTGAA ATTCTAATAG GAGGACACTA GAAAATGGCA TTAAAGACA ATGCAGTTGA	1440
50	ATTAGAAGAA CGCGTAGTTG CTGTCAACCG TGTTACAAA GTTGTTAAAG GTGGACGTCG	1500
	TCTTCGTTT GCAGCTCTTG TTGTTGTTGG TGACCACAAT GGTCGCGTAG GATTTGGTAC	1560
55	TGGTAAAGCT CAAGAAGTTC CAGAAGCAAT CCGTAAAGCA GTAGATGATG CTAAGAAAAA	1620
	CTTGATCGAA GTTCCTATGG TTGGAACAAC AATCCACAC GAAGTTCTTT CAGAATTCGG	1680
	TGGAGCTAAA GTATTGTTGA AACCTGCTGT AGAAGGTTCT GGAGTTGCCG CTGGTGGTGC	1740
60	AGTTCGTGCC GTTGTGGAAT TGGCAGGTGT GGCAGATATT ACATCTAAAT CACTTGGTTC	1800

-79-

	TAACACTCCA ATCAACATTG TTCGTGCAAC TGTGAAGGT TTGAAACAAT TGAAACGCGC	1860
	TGAAGAAATT GCTGCCCTTC GTGGTATTTT AGTTTCTGAT TTGGCATAAG AAAGGGGATA	1920
5	AAATGGCTCA AATTAATAATT ACTTTGACTA AGTCTCCAAT CGGACGCATT CCATCACAAC	1980
	GTAAAACTGT TGTAGCACTT GGAAGTGGCA AATTGAACAG CTCTGTTATT AAAGAAGATA	2040
10	ACGCTGCTAT CCGTGGTATG ATTACAGCAG TATCTCACTT AGTAACAGTT GAAGAAGTAA	2100
	ACTAATGAAG TTTTAGGGGA TGTGCACTGT ACCATCCCCT AAAACTAGAT ATAGTCATCT	2160
	ATGATGACAT CGTATAGGCG AGTTGATGGG GGAGACAACC TTTCTCCCT TATCGGCGCT	2220
15	AGCATTTTAC AAAAGAGGAG AAAATAAAAA TGAACTTCA TGAATTGAAA CCTGCAGAAG	2280
	GTTCTCGTAA AGTACGTAAC CGCGTTGGTC GTGGTACTTC ATCAGGTAAC GGTAACACAT	2340
20	CTGGTCGTGG TCAAAAAGGT CAAAAGCTC GTAGCGGTGG CGGAGTTCGC CTTGGTTTTG	2400
	AAGGTGGACA AACTCCATTG TTCCGTCGTC TTCCAAAACG TGGATTCACT AACATCAACG	2460
	CTAAGAATA CGCAATTGTG AACCTTGACC AATTGAACGT CTTGAAGAT GGTGCTGAAG	2520
25	TAATCCAGT TGTCTTATC GAAGCAGGAA TTGTTAAAGC TGAAAAGTCA GGTATTAAAA	2580
	TTCTTGGTAA CGGTGAGTTG ACTAAGAAAT TGAAGTGAA AGCAGCTAAA TTCTCTAAAT	2640
30	CAGCTGAAGA AGCTATCACT GCTAAAGGTG GTTCAGTAGA AGTCATCTAA GAGAGGTGAC	2700
	CTATGTTTTT TAAATTATTA AGAGAAGCTC TTAAAGTCAA GCAGGTTCGA TCAAAAATTT	2760
	TATTTACAAT TTTTATCGTT TTGGTCTTTC GTATCGGAAC TAGCATTACA GTTCTGGTG	2820
35	TGAATGCCAA TAGCTTGAAT GCTTTAAGTG GATTATCCTT CTAAACATG TTGAGCTTGG	2880
	TGTCGGGGAA TGCCCTAAA AACTTTTCGA TTTTGGCCCT AGGAGTTAGT CCCTATATCA	2940
40	CCGCTTCTAT TGTGTCCAA CTCTTGCAA TGGATATTTT ACCCAAGTTT GTAGAGTGGG	3000
	GTAACAAGG GGAAGTAGGT CGAAGAAAAT TGAATCAAGC TACTCGTTAT ATTGCTCTAG	3060
	TTCTCGCTTT TGTGCAATCT ATCGGGATTA CAGCTGGTTT TAATACCTTG GCTGGAGCTC	3120
45	AATTGATTAA AACTGCTTTA ACTCCACAAG TTTTCTGAC GATTGGTATC ATCTTAACAG	3180
	CTGGTAGTAT GATTGTCACT TGGTTGGGTG AGCAAATTAC AGATAAGGA TACGGAAACG	3240
50	GTGTTTCCAT GATTATCTTT GCCGGGATTG TTTCTCAAT TCCAGAGATG ATTCAGGGCA	3300
	TCTATGTGGA CTACTTTGTG AACGTCCCAA GTAGCCGTAT CACTTCATCT ATCATTTTCG	3360
	TAATCATTTT GATTATTACT GTATTGTTGA TTATTACTT TACAACTTAT GTTCAACAAG	3420
55	CAGAATACAA AATTCCAATC CAATATACTA AGGTGACACA AGGTGCTCCA TCTAGCTCTT	3480
	ACCTCCGTT AAAGGTAAAT CCTGCTGGAG TTATCCCTGT TATCTTGCC AGTTCGATTA	3540
60	CTGCAGCGCC TGCGGCTATT CTCAGTTTT TGAGTGCCAC AGGTCATGAT TGGGCTTGGG	3600
	TAAGGGTAGC ACAAGAGATG TTGGCAACTA CTTCTCCAAC TGGTATTGCC ATGTATGCTT	3660

-80-

	TGTTGATTAT TCTCTTTACA TTCTTCTATA CGTTTGTACA GATTAATCCT GAAAAAGCAG	3720
5	CAGAGAGCCT ACAAAGAGT GGTGCCTATA TCCATGGAGT TCGTCCTGGT AAAGGTACAG	3780
	AAGAATATAT GTCTAAACTT CTTCTGTCGTC TTGCAACTGT TGGTTCCTC TTCCTTGGTG	3840
	TGATTTCCAT TTTACCGATT GCAGCTAAAG ATGTATTTGG TCTTTCTGAT GTTGTGTCCT	3900
10	TTGGTGGAAC AAGTCTCTTG ATCATTATCT CTACAGGTAT CGAAGGAATC AAGCAATTGG	3960
	AAGGTTACCT ATTGAAACGT AAGTATGTTG GTTTCATGGA CAGAACAGAA TAAAAGTATT	4020
	TACTGAATCA GTAAATACTG AGGGAGTGGA GGTTTAACT CTGACATTG TAAGAGTTGG	4080
15	ATCTCCCCTC TTCTATTTTG TTTTAAATC GGGGTGAAAA AACTTTTTGC TTCTATTTAA	4140
	AAACAAAATA AGGAGATCAA ATCATGAATC TTTTGATTAT GGGCTTACCT GGTGCAGGTA	4200
20	AGGGAECTCA AGCAGCAAAA ATCGTAGAAC AATTCCATGT TGCACATATC TCAACAGGTG	4260
	ATATGTTCCG CGCTGCAATG GCAAATCAAA CTGAAATGGG TGTTCTTGCT AAGTCATATA	4320
	TTGACAAGGG TGAATTGGTT CCTGACGAAG TTACAAATGG AATCGTAAAA GAACGCCTTT	4380
25	CACAAGATGA TATTAAAGAA ACAGGATTCT TATTGGATGG TTACCCACGT ACAATTGAAC	4440
	AAGCTCATGC CTTGGACAAA ACATTGGCTG AACTTGGCAT TGAAGTAGAA GGTATTATCA	4500
30	ATATTGAAGT GAACCCTGAC AGCCTCTTGG AACGTTTGAG TGGCCGTATC ATCCACCGCG	4560
	TAAGTGGAGA AACTTTCCAC AAGGTCTTTA ACCCACCAGT TGACTATAAA GAAGAAGATT	4620
	ACTACCAACG TGAAGATGAT AAGCCTGAGA CAGTAAAACG TCGTTTGGAT GTTAATATTG	4680
35	CTCAAGGAGA ACCAATCATT GCTCACTACC GTGCCAAAGG TTTGGTTCAT GACATCGAAG	4740
	GTAATCAAGA TATCAATGAT GTCTTCTCAG ATATTGAAA AGTATTGACA AATTTGAAAT	4800
40	AAAGCGTTT TCACACTTGC AAAAATCCGC TACAAATGTT ATACTGAAAT AGTCTGACTT	4860
	ATAATTGTTG TCTCTGTGTC TAGAGGCATC GAATCGAAAT TTATGGAGGT GCTTTTGCCT	4920
	GGCAAAAGAC GATGTGATTG AAGTTGAAGG CAAAGTAGTT GATACAATGC CGAATGCAAT	4980
45	GTTTACGGTT GAACTTGAAA ATGGACATCA GATTTTAGCA ACAGTTTCTG GTAAAATTCG	5040
	TAAAACTAT ATTCTGATTT TAGCGGGAGA TCGTGTTACT GTCGAAATGA GTCCATATGA	5100
50	CTTGACACGT GGACGTATCA CTTACCGCTT TAAATAATCG AAAAAGTTG AGGGATAAGA	5160
	AATGAAAGTA AGACCATCGG TCAAACCAAT TTGCGAATAC TGTAAGTTA TTCGTCGTAA	5220
	TGGTCGTGTT ATGGTAATTT GCCCAGCAAA TCCAAAACAC AAACAACGTC AAGGATAAGA	5280
55	TAGAAAGGAG AAAACATGGC TCGTATTGCT GGAGTTGATA TTCCAAATGA CAAACGCGTA	5340
	GTAATCTCAT TGACTTATGT TTATGGTATC GGACTTGCAA CATCTAAGAA AATTTTGGCT	5400
60	GCTGCTGGAA TCTCAGAAGA TGTTCTGTGA CGTGATCTTA CATCAGATCA AGAAGATGCT	5460

-81-

ATCCGTCGTG AAGTGGATGC AATCAAAGTT GAAGGTGACC TTCGTCGTGA AGTAACTTG 5520  
AACATCAAAC GTTTGATGGA AATCGGTTCA TACCGTGGTA TCCGTCACCG TCGTGGACTT 5580  
5 CCTGTCCGTG GACAAAATAC TAAAAACAAC GCTCGCACTC GTAAAGGTAA AGCTGTTGCG 5640  
ATTGCTGGTA AGAAAAATA ATATAGGAGG TAAAAGTCTT GGCTAAACCA ACACGTAAAC 5700  
GTCGTGTGAA AAAGAATATC GAATCTGGTA TTGCTCATAT TCACGCTACA TTTAATAACA 5760  
10 CTATTGTTAT GATTACTGAT GTGCATGGTA ATGCAATTGC TTGGTCATCA GCTGGTGCTC 5820  
TTGGTTTCAA AGGTTCTCGT AAATCTACAC CATTGCTGCTC TCAAATGGCT TCTGAAGCTG 5880  
15 CTGCTAAATC TGCACAAGAA CACGGTCTTA AATCAGTTGA AGTTACTGTA AAAGGTCCAG 5940  
GTTCTGGTCG TGAGTCAGCT ATTCGTGCGC TTGCTGCCGC TGGTCTTGAA GTAACAGCAA 6000  
TTCGTGATGT GACTCCAGTG CCACACAATG GTGCTCGTCC TCCAAAACGT CGCCGTGTAT 6060  
20 AATCATCGCA TTACTGCT TTTCTTTAA GAGGGAGTAA CTAAATGATC GAGTTTGAAA 6120  
AACCAAATAT AACAAAAATT GATGAAAATA AAGATTATGG CAAGTTTGTA ATCGAACCAC 6180  
25 TTGAACGTGG CTACGGTACA ACTCTTGGA ACTCTCTTCG TCGTGTACTT CTAGCTTCTC 6240  
TACCAGGAGC AGCTGTGACA TCTATCAACA TTGATGGTGT GTTACATGAG TTTGACACAG 6300  
TTCCAGGTGT TCGTGAAGAC GTGATGCAAA TCATTCTGAA CATTAAAGGA ATTGCAGTGA 6360  
30 AATCGTACGT TGAAGACGAA AAAATCATCG AACTGGATGT TGAAGGTCCT GCTGAAGTAA 6420  
CAGCTGGTGA CATTTTGACA GATAGCGATA TTGAAATTGT AAATCCAGAT CATTATCTCT 6480  
35 TTACAATCGG TGAAGGTTCT TCTCTAAAAG CGACTATGAC TGTTAACAGT GGTCGTGGAT 6540  
ATGTACCTGC TGACGAAAAT AAAAAGGATA ATGCACCACT TGGAACACTT GCTGTAGATT 6600  
CTATTTATAC ACCAGTTACA AAAGTCAACT ATCAAGTGGA ACCTGCTCGT GTAGGTAGCA 6660  
40 ATGATGGTTT CGACTCTAG 6679

## (2) INFORMATION FOR SEQ ID NO:34:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1703 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
50  
(ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
55 (iv) ANTI-SENSE: NO  
60  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

-82-

	AGAATACCTT GGGGCAACTG TTCAAGTCAT TCCTCATATC ACAGATGCTT TGAAAGAAAA	60
	AATCAAGAGT GCCGCTCTAA CGACCGACTC TGATGTCATT ATCACAGAGG TTGGTGGAAAC	120
5	AGTAGGAGAT ATCGAGTCCT TGCCATTTCCT AGAGGCTCTT CGCAGATGAA GGCAGATGTG	180
	GGGCGGATAA TGTATGTAT ATCCATACAA CCTTCTTCCT TACCTCAAGG CTGCTGGTGA	240
	AATGAAACCA AACCAACCCA ACACTCTGTC AAAGATTGCG TGGCTTGGGA ATCCAACCAA	300
10	ATATGTTGGT TATTCGTACA GAAGAGCCAG CTGGTCAAGG AATTAAAAAT AACTGGCCCC	360
	AGTTCTGTGA TGTGGCACCA GAATCCCTAA TCGAATCGTT GGATGTTGAA CACCTTTACC	420
15	AAATTCCACT GAACTTGCG AGACAAGGGA TGGACCAAAT TGTTTGTGAT CATTTGAAAT	480
	TAGACGCACC AGCAGCGGAT ATGACAGAAT GGTGAGCCAT GGTGGACAAG GTCATGAACC	540
	TCAAGAAACA AGTTAAGATT TCCCTTGTTG GTAAGTATGT GGAGTTGCAA GATGCCTATA	600
20	TCTCAGTGGT CGAAGCCTTG AAACACTCTG GCTATGTCAA TGATGTAGAA GTTAAATCA	660
	ATTGGGTCAA TGCCAATGAT GTGACAGCAG AGAATGTAGC AGAACTCTTG TCTGATGCGG	720
25	ACGGGATCAT CGTACCAGGT GGTTTTGGTC AACGTGGTAC AGAAGGGAAA ATCCAAGCCA	780
	TCCGCTATGC GCGTGAAAAT GATGTTCCAA TGTTGGGAGT CTGCTTGGGA ATGCAGTTGA	840
	CATGTATCGA GTTTGCTCGT CACGTTTTAG GTCTTGAAGG TGCCAATTCT GCAGAGCTTG	900
30	CACCAGAAAC AAAATACCCT ATCATTGATA TCATGCGTGA TCAGATTGAT ATTGAGGATA	960
	TGGGTGGAAAC CCTTCGTTTG GGACTTTATC CGTCTAAGTT GAAACGTGGC TCTAAGGCTG	1020
35	CTGCTGCTTA TCACAATCAA GAAGTGGTGC AACGCCGTCA CCGTCACCGT TATGAGTTTA	1080
	AATAATGCCT TCCGTGAGCA GTTTGAGGCA GCAGGTTTGT CTTTTCAGGA GTTTCTCCAG	1140
	ACAATCGTTT GGTAGAAATC GTGGAAATCC TGAAAATAAA TTCTTTGTAG CTTGTCAGTA	1200
40	TCACCCTGAA CTGTCAGCCG TCCAACCGAC CAGAAGAACT CTACACTGCC TTTGTTACTG	1260
	CAGCGGTTGA GAACAGCAAT TAGCAAAATC AGAACCTTTG AGAAAAATCT CAGAGGTTTT	1320
45	TTGCATACGA TGATATTGCA GTATATCTGA GGTAGGAGTC CTCTGTATGT ACCTGCTACC	1380
	GTTGAAATCA ATAGCGACTC CCTCTTGCCC TGTGCTAGTG AATGGATTTA TCAGTATATT	1440
	GAAATGAAAT AAAATTTGAA CAAATTAATT CGGAAAGCCA AATCAATTTC TAGCAAAGTT	1500
50	TTAGGAACTG GATTGTATAG TGAATTGAAA TAAGATGTGA ACATCTCTAT CAGGAAAGTC	1560
	AAATTAATTT ATAGAAATAT TTTAGCAGTC AAGATGGACT GTTATAGATT CAATATACTA	1620
55	TACTTTTTTA ATTTAATCCA CTATAATAAA ATGAAATAAT AACAGGACAA ATCGTTCAGG	1680
	ACAGTCAAAT CGACTCTAGA GGA	1703

60 (2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

-83-

(A) LENGTH: 1620 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ATTGTAAAAC	ACCAAGGAAA	AACAGCTAAA	GAAGCGAAAG	AATTGGCCAT	TGACTACATG	60
AATAAGGTTG	GCATTCCAGA	CGCAGATAGA	CGTTTTAATG	AATACCCATT	CCAATATTCT	120
GGAGGAATGC	GTCAACGTAT	CGTTATTGCG	ATTGCCCTTG	CCTGCCGACC	TGATGTCTTG	180
ATCTGTGATG	AGCCAACAAC	TGCCTTGGAT	GTAACATTC	AAGCTCAGAT	TATTGATTTG	240
CTAAAATCTT	TACAAAACGA	GTATCATTTT	ACAACAATCT	TTATTACCCA	CGACCTTGGT	300
GTGGTGGCAA	GTATTGCGGA	TAAGGTAGCG	GTTATGTATG	CAGGAGAAAT	CGTTGAGTAT	360
GGAACGGTTG	AGGAAGTCTT	CTATGACCCT	CGCCATCCAT	ATACATGGAG	TCTCTGTCTT	420
AGCTTGCCCT	AGCTTGCTGA	TGATAAAGGG	GATCTTTACT	CAATCCCAGG	AACACCTCCG	480
TCACTTTATA	CTGACCTGAA	AGGGGATGCT	TTTGCCTTGC	GTTCTGACTA	CGCAATGCAG	540
ATTGACTTCG	AACAAAAAGC	TCCTCAATTC	TCAGTATCAG	AGACACATTG	GGCTAAAAC	600
TGGCTTCTTC	ATGAGGATGC	TCCAAAAGTA	GAAAAACCAG	CTGTGATTGC	AAATCTCCAT	660
GATAAGATCC	GTGAAAAAAT	GGGATTTGCC	CATCTGGCTG	ACTAGGAGGA	AGGAAATGTC	720
TGAAAAATTA	GTAGAAATCA	AAGATTTAGA	AATTCCTTC	GGTGAAGGAA	GTAAGAAGTT	780
TGTCGCGGTT	AAAAATGCTA	ACTTCTTTAT	CAACAAGGGA	GAAACTTTCT	CGCTTGTAGG	840
TGAGTCCGGT	AGTGGGAAAA	CAACTATTGG	TCGTGCTATC	ATCGGTCTAA	ATGATACAAG	900
TAATGGAGAT	ATCATTTTGT	ATGGTCAAAA	GATTAATGGT	AAGAAATCGC	GTGAACAAGC	960
TGCGGAATTG	ATTCGTCGAA	TCCAGATGAT	TTTCCAAGAC	CCTGCCGCAA	GTTTGAATGA	1020
ACGTGCGACT	GTTGATTATA	TTATTCTGA	AGGTCTTTAC	AATCACCATT	TATTTAAGGA	1080
TGAAGAAGAA	CGTAAAGAGA	AAGTTCAAAG	TATTATCCGT	GAAGTAGGTC	TTCTTGCTGA	1140
GCACTTGACT	CGTTACCCTC	ATGAATTCTC	AGGCGGTCAA	CGTCAACGTA	TCGGTATTGC	1200
CCGTGCCTTG	GTCATGCAAC	CAGACTTTGT	TATTGCAGAT	GAGCCAATTT	CAGCCTTGGA	1260
CGTTTCTGTA	CGTGCCCAAG	TCTTGAACCT	GCTCAAAAAA	TTCCAAAAAG	AGCTCGGCTT	1320
GACCTATCTC	TTCATCGCCC	ATGACTTGTC	GGTTGTTGCG	TTTATTTTCT	ATCGTATCGC	1380



-84-

AGTTATTTAC AAGGGTGTTA TTGTAGAGGT TGCAGAAACA GAAGAATTGT TTAACAATCC 1440  
AATTCACCCA TATACTCAAG CCTTGCTTTC AGCGGTACCA ATCCCAGATC CAATCTTGGA 1500  
5 ACGTAAGAAG GTCTTGAAGG TTTACGACCC AAGTCAACAC GACTATGAGA CTGATAAGCC 1560  
ATCTATGGTA GAAATCCGTC CAGGTCACCTA TGTTCGGGCG AACCAAGCCG AATTAGCACG 1620

10 (2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:  
15 (A) LENGTH: 984 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)  
20 (iii) HYPOTHETICAL: NO  
(iv) ANTI-SENSE: NO

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

30 GTACCCGGGG ATCAGGTTTT ACGGATTCTT GAAGTTCTCT GTGGGCAGGA CCTCTTGCG 60  
GTAAGAGTAA GAGTGATTCT ACAAGATTTA CTAGAAGCTA GAAAAATGTG GCAAGCTAAT 120  
GTCAGCTTTC AAAATGCCAT GGAATATCTG GTCTTGAAAG AAATATAAAC TCAAAAATGA 180  
35 ATGATAAGA AAGGAAAGGG CTGTTTTATG GACAAAAAG AATTATTTGA CGCGCTGGAT 240  
GATTTTTCCC AACAGTTATT GGTAACCTTG GCCGATGTGG AAGCCATCAA GAAAAATCTC 300  
40 AAGAGCCTGG TAGAGGAAAA TACAGCTCTT CGTTTGAAA ATTCTAAGTT GCGAGAACGC 360  
TTGGGTGAGG TGGAAGCAGA TGCTCCTGTC AAGGCCAAGC ATGTTCTGTA AAGTGTCCTG 420  
CGCATTACC GTGATGGATT TCACGTATGT AATGATTTTT ATGGACAACG TCGAGAGCAG 480  
45 GACGAGGAAT GTATGTTTTG TGACGAGTTG CTATACAGGG AGTAGGCATG CAGATTCAA 540  
AAAGTTTTAA GGGGCAGTCT CCCTATGGCA AGCTGTATCT AGTGGCAACG CCGATTGGCA 600  
50 ATCTAGATGA TATGACCTTT CGAGCTATCC AGACCTTGAA AGAAGTAGAT TGGATTGCTG 660  
CTGAGGATAC GCGCAATACA GGTCTTTTGC TCAAGCATTT TGACATTTCC ACCAAGCAGA 720  
TCAGTTTTCA TGAGCACAAT GCCAAGGAAA AAATTCCTGA TTTGATTGGT TTCTTGAAAG 780  
55 CAGGGCAAAG TATTGCTCAG GTCTCTGATG CCGGTTTGCC TAGCATTTCA GACCCTGGTC 840  
ATGATTTAGT TAAGGCAGCT ATTGAGGAAG AAATTGCAGT TGTGACAGTT CCAGGTGCCT 900  
60 CTGCAGGAAT TTCTGCCTTG ATTGCCAGTG GTTTAGCGCC ACAGCCACAT ATCTTTTACG 960

-85-

GTTTTTTACC GAGAAAATCA GGTC

984

(2) INFORMATION FOR SEQ ID NO:37:

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1554 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CTAGAGTCGA AAAGACAAGC AGGAGCGTAT TTCCAAAGAA ACCATGGAAA TCTATGCCCC	60
GCTTGCCCAT CGTTTGGGGA TTTCCAGTGT CAAATGGGAA TTAGAAGACT TGTCTTTCCG	120
25 TTATCTCAAT CCAACGGAGT TTTACAAGAT TACCCATATG ATGAAGGAAA AGCGCAGGGA	180
GCGTGAGGCC TTGGTGGATG AGGTAGTCAC AAAATTAGAG GAGTATACGA CAGAACGTCA	240
30 CTTGAAAGGG AAGATTTATG GTCGTCCCAA GCATATTTAC TCAATTTTCC GCAAAATGCA	300
GGACAAGAGA AAACGGTTTG AGGAAATCTA TGATCTGATT GCTATTCGTT GTATTTTAGA	360
TACCCAAAGT GATGTTTATG CCATGCTTGG TTACGTGCAT GAATTTTGA AACCGATGCC	420
35 AGGTCGCTTC AAAGACTATA TCGCCAACCG CAAGGCCAAT GGTATCAGT CTATCCATAC	480
GACTGTTTAT GGACCAAAG GGCCGATTGA ATTCCAGATT CGAACCAAGG AAATGCACGA	540
40 GGTGGCTGAG TACGGGGTTG CGGCTCACTG GGCTTATAAG AAAGGTATAA AGGGGCAAGT	600
TAACAGCAAG GAATCAGCTA TTGGAATGAA CTGGATCAAG GAGATGATGG AGCTCCAAGA	660
CCAGGCTGAT GATGCTAAGG AATTTGTGGA CTCTGTTAAG GAAAACTATT TGGCTGAGGA	720
45 GATTACCGTT TTACCCAGA TGGAGCTGTC CGTTCCTTCC CAAAGATTCA GGACCGATTG	780
ATTTTGCTTA CGAAATCCAT ACCAAGGTCG GTGAAAAGCA ACTGGTGCCA AGGTCAATGG	840
50 CCGCATGGTT CCACTGACAC CCAAGTTAAA GGACAGGGGA TCAGGTTGAA ATTATCGCCA	900
ACCCGAATC CTTTGGACCT TAGCCGTGAC TGGCTCAATA TGGTCAAGAC TAGCAAGGCG	960
CGCAATAAGA TTCGCCAGTT CTTTAAAAAC CAAGATAAGG AATTGTCTGT CAACAAGGGT	1020
55 CGTGAGATGC TGATGGCTCA GTTCCAAGAA AATGGCTATG TGGCAAATAA ATTTATGGAC	1080
AAGCGCCACA TGGATCAAGT TCTGCAAAAG ACCAGTTACA AGACAGAAGA CTCCCTCTTT	1140
60 GCGGCCATTG GTTTTGGGGA AATCGGTGCG ATTACCGTCT TTAACCGTCT GACTGAAAAG	1200

-86-

GAGCGCCGTG AGGAAGAGCG TGCCAAGGCC AAGGCTGAGG CAGAGGAGCT TGTCAAAGGT 1260  
GGCGAGGTCA AGGTTGAAAA TAAAGAACT CTCAAGGTCA AGCATGAGGG GGGAGTGGTT 1320  
5 ATTGAAGGTG CTTCTGGTCT CCTAGTGCGG ATTGCTAAGT GTTGTAAACC CGTGCCTGGT 1380  
GACGATATTG TTGGCTACAT TACCAAGGGT CGTGGTGTGG CTATTCACCG TGTGGACTGT 1440  
ATGAACCTGC TGCCCAAGA AACTACGAG CAACGTCTCC TTGATGTGGA ATGGGAAGAC 1500  
10 CAGTACTCTA GCTCAAATAA GGAGTATATG GCCCATATCG ACTCTAGAGG ATCC 1554

## (2) INFORMATION FOR SEQ ID NO:38:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3190 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CGTCAGTGCT AAAACAGGGG AAATTCTGGC AACAACGCAA CGACCGACCT TTGATGCAGA 60  
TACAAAAGAA GGCATTACAG AGGACTTTTT TGGCGTGATA TCCTTTACCA AAGTAACTAT 120  
35 GAGCCAGGTT CCACTATGAA AGTGATGATG TTGGCTGCTG CTATTGATAA TAATACCTTT 180  
CCAGGAGGAG AAGTCTTTAA TAGTAGTGAG TTAAAAATTG CAGATGCCAC GATTCGAGAT 240  
40 TGGGACGTTA ATGAAGGATT GACTGGTGGC AGAATGATGT CTTTTTCTCA AGGTTTTGCA 300  
CACTCAAGTA ACGTTGGGAT GACCCTCCTT GAGCAAAAGA TGGGAGATGC TACCTGGCTT 360  
GATTATCTTA ATCGTTTTAA ATTTGGTGTT CCGACCCGTT TCGGTTTGAC GGATGAGTAT 420  
45 GCTGGTCAGC TTCCTGCGGA TAATATTGTC AACATTGCGC AAAGCTCATT TGGACAAGGG 480  
ATTTCACTGA CCCAGACGCA AATGATTTCG GCCTTTACAG CTATTGCTAA TGACGGTGTC 540  
50 ATGCTGGAGC CTAAATTTAT TAGTGCCATT TATGATCCAA ATGATCAAAC TGCTCGGAAA 600  
TCTCAAAAAG AAATTGTGGG AAATCCTGTT TCTAAAGATG CAGCTAGTCT AACTCGGACT 660  
AACATGGTTT TGGTAGGGAC GGATCCGGTT TATGGAACCA TGTATAACCA CAGCACAGGC 720  
55 AAGCCAACTG TAACTGTTCC TGGGCAAAAT GTAGCCCTCA AGTCTGGTAC GGCTCAGATT 780  
GCTGACGAGA AAAATGGTGG TTATCTAGTC GGGTTAACCG ACTATATTTT CTCGGCTGTT 840  
60 CGATGAGTCC GGCTGAAAAT CCTGGATTTT ATCTTGATG TGACGGTCCA ACAACCTGGA 900

- 87 -

	ACATTATTCA GGTATTCAGT TGGGAGAATT TGCCAATCCT ATCTTGGAGC GGGCTTCAGC	960
	TATGAAAGAC TCTCTCAATC TTCAAACAAC AGCTAAGGCT TTGGAGCAAG TAAGTCAACA	1020
5	AAGTCCTTAT CCTATGCCCA GTGTCAAGGA TATTTACCT GGTGATTTAG CAGAAGAATT	1080
	GCGTCGCAAT CTTGTACAAC CCATCGTTGT GGGAACAGGA ACGAAGATTA AAAACAGTTC	1140
10	TGCTGAAGAA GGGAGAATC TTGCCCCGAA CCAGCAAGTC CTTATCTTAT CTGATAAAGC	1200
	AGAGGAGGTT CCAGATATGT ATGGTTGGAC AAAGGAGACT GCTGAGACCC TTGCTAAGTG	1260
	GCTCAATATA GAACTTGAAT TTCAAGGCTC GGGCTCTACT GTGCAGAAGC AAGATGTTCTG	1320
15	TGCTAACACA GCTATCAAGG ACATTAAAAA AATTACATTA ACTTTAGGAG ACTAATATGT	1380
	TTATTTCCAT CAGTGCTGGA ATTGTGACAT TTTTACTAAC TTTAGTAGGA ATTCCGGCCT	1440
20	TTATCCAATT TTATAGAAAG GCGCAAATTA CAGGCCAGCA GATGCATGAG GATGTCAAAC	1500
	AGCATCAGGC AAAAGCTGGG ATTCCTACAA TGGGAGGTTT GGTTTTCTTG ATTACTTCTG	1560
	TTTTGGTTGC TTTCTTTTTC GCCCTATTTA GTAGCCAATT CAGCAATAAT GTGGGAATGA	1620
25	TTTTGTTCAT CTTGGTCTTG TATGGCTTGG TCGGATTTTT AGATGACTTT CTCAAGGTCT	1680
	TTCGTAAAAT CAATGAGGGG CTTAATCCTA AGCAAAAATT AGCTCTTCAG CTTCTAGGTG	1740
30	GAGTTATCTT CTATCTTTTC TATGAGCGCG GTGGCGATAT CCTGTCTGTC TTTGGTTATC	1800
	CAGTTCATTT GGGATTTTTC TATATTTTCT TCGCTCTTTT CTGGCTAGTC GGTTTTTCAA	1860
	ACGCAGTAAA CTGACAGAC GGTGTTGTAC GGTTTAGCTA GTATTTCCGT TGTGATTAGT	1920
35	TTGTTTGCCT ATGGAGTTAT TGCCTATGTG CAAGGTCAGA TGGATATTCT TCTAGTGATT	1980
	CTTGCCATGA TTGGTGGTTT GCTCGGTTTC TTCATCTTTA ACCATAAGCC TGCCAAGGTC	2040
40	TTTATGGGTG ATGTGGGAAG TTTGGCCCTA GGTGGGATGC TGGCAGCTAT CTCTATGGCT	2100
	CTCCACCAGG AATGGACTCT CTTGATTATC GGAATGTGT ATGTTTTTGA AACAACTTCT	2160
	GTTATGATGC AAGTCAGTTA TTTCAAACG ACAGGTGGTA AACGTATTTT CCGTATGACG	2220
45	CCTGTACATC ACCATTTTGA GCTTGGGGGA TTGTCTGGTA AAGGAAATCC TTGGAGCGAG	2280
	TGGAAGGTTG ACTTCTTCTT TTGGGGAGTT GGGCTTCTAG CAAGTCTCCT GACCCTCGCA	2340
50	ATTTTGTATT TGATGTAAGA ATGGCACCCT GATGTTTCAG GGTGTTTTTG TGTTTAAATA	2400
	CACAATGAAA ATCAAAGAAC AAAC TAGAAA GCTAACTTTA GGCTGCTCAA AATATAATAT	2460
	ATTGAAACTA GAATAGTACA CCTCTACTTC TAAACATTG TTAGAAATCG ATTTGACTGT	2520
55	CCTGAACGAT TTATCCTGTT CTTATTTTCAT TTTACTATAC AGTTTCGAGG TTGTAGATAA	2580
	GGCGAAGCTG ATGTGGTTTG AAGAGATTTT CTGAAAAGTG TTAACACCTA CAGACAAGCC	2640
60	TGACGATAGC AAGAACTACC CTA CTGATA GGTATCGGCT TTTGCTTTCT GAAAAAATT	2700
	ATTTTAAGCA TTTGACAAAT CTAGCAACAA AAAATTCTAT AAATATAATA GATTGAACT	2760

-88-

AGAATAGTAC ACATCTACTT CTAAACATT GTTAGAAATC GATTTGACTG TCCTGATCGA 2820  
TTTGTCTGT TCTTGTTTCA TTTTACTATA TTTCTATGAT AAAACGCATA GTATCAAGTT 2880  
5 TTCTTAATCC CCTGATACTA TGCGTGTGTTG TAATTTTTTAA GATTTTGTGC TTAGAGTCGA 2940  
CTCCTTATTT TAGATATTTA AAAGGAATCT CACTTCCACA GAGCCAGTTG TAGACTTGGT 3000  
10 CATTAACAAA TACATTCATG GCTTCGTGAG CATACTCAGG CATGATACGA TAGGTTTTAT 3060  
CGCAGGTCAG ACGATTATAA ATCGCAAAC GGGTAATGGG ATAGCAAACA TCGTCGTCCA 3120  
AGCCCGTAAT CATCTTAACC TCACCTTGA TACGATGGGC AAGATTTTTG ACATCGACTC 3180  
15 TAGAGGATCC 3190

## (2) INFORMATION FOR SEQ ID NO:39:

20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5992 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

30 (iv) ANTI-SENSE: NO

## 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TTGTTCTTAG TGTTCCGACA AAGATTCTTC AAAATCAAAT CATGGAAGAA GAAGGTAAAC 60  
GTCTCAAGGA AGTGTTCCAT ACAGATATTC ATAGCTTAAA GGGACCACAA AATTATCTGA 120  
40 AGTTGGATGC CTTTTATCAT TCTTGCAGGA AAATGATGAA AATCGCTTAT TTAGACGCTT 180  
TAAAATGCAA GTCTTGGTCT GGCTTACTGA GACAGAGACA GGAGATTGG ATGAAATCGG 240  
45 GCAACTCTAC CGTTACCAAC ATTTTCTAGC AGACCTTCGT CATAATGGGA ATTTATCATC 300  
CCAGAGCTTA TTTGTGACGG AAGATTTTTG GAAACGTAGT CAAGAAAGGG CAGAGACTTG 360  
CAAGCTTTTA GTGACTAATC ATGCCTATCT CGTAACCAGA CTTGAAGATA ATCCTGAATT 420  
50 TGTCAGTGAC CGTTTACTGA TTATTGATGA AGTCCAAAAG ATTTTGTTAG CTCTAGAAAA 480  
TCTGCTTCAA GAGACCTACG ATATACAATC TATTATCGAT TTAATTGATA AGGCTTTAGT 540  
55 AGGAGAAGAA AACAGGGTTC AACACGGAT ACTAGAAAGT ATTCGCTTTG AGTGCCTCTA 600  
CTTGATAGAA CAATTTCACT CTGGCAAATC TAGGAAAAAT ATCTTAGATT CTCTGGACAA 660  
TCTCCATCAG TATTTTTCAG AATTAGAAGT GGAAGGCTTT GATGAGCTGG TTCGCTATTT 720  
60 TACAGCTGAA GGTGATTACT GGCTTGAAGT AACTGAAACG AGTCAAAGA AAATTCAGAT 780

	TTCTTCTACA	AAATCAGGCC	GTA CTCTCT	GTCCTCTTTA	CTTCCTGAGA	GTTGCCAAGT	840
5	CTTGGGAGTA	TCGGCTACTC	TTGAGATTAG	TCAGAGGGTT	TCTTTGGCAG	ACCTTTTAGG	900
	CTATCCTGAA	GCCAAATTTG	TCAAGATTGA	ATCTCGGGGA	AAACAGGAAC	AAGAAGTGGT	960
	TATGGTCAAA	GATTTCCCTC	TGGTAACAGA	AACCTCCTTA	GAAGTCTATG	CCAGAGAGGT	1020
10	AGCTGCTTTA	CTAGTGGA	TTCAAGCTTT	CCAGCAACCG	ATTTTGGTTC	TCTTTACCGC	1080
	TAAAGACATG	CTTCTAGCAG	TATCGGATTT	ACTTACAGTT	AGCCACTTGG	CCCAGTATAA	1140
15	AAATGGGGAT	GTTTCATCAGC	TAAAGAAACG	CTTTGAAAAA	GGTGAACAAC	AAATCTTGCT	1200
	TGGTGCAGCA	AGTTTCTGGG	AGGGAGTTGA	TTTTTCAAGC	CATCCTTTTG	TGATTCAAGT	1260
	TGTACCGAGG	CTTCCTTTCC	AAAATCCTCA	AGAACCCTTG	ACGAAAAAGA	TTAATCAAGA	1320
20	ACTGAATCAA	GAAGGGAAAA	ATGCCTTTTA	TGATTATCAA	TTGCCAATGG	CCATTATTCTG	1380
	TTTAAACAG	GCTTTGGGAA	GAAGTATGAG	ACGTGAATAC	CAACGTTTCT	TAACCTTTAT	1440
25	TTTGGATAGG	AGAATCATCG	GAAAACGATA	CGGCAAACAA	ATAGTAGCAT	CTCTAGCAGA	1500
	AGAAGCGACT	GTTAAACCA	TCTCTCGATC	CGAAGTTGAC	GAGGCTATTG	ATAGATTTTT	1560
	TAATGAACTT	TGATAAATAG	TATTGTATGA	AAGTATAAGG	TTAGTACATA	TGAAACGTTT	1620
30	TCTCGACTCT	AGAGTCGATT	ATAGTTTGCT	CTTGCCAGTA	TTTTTTCTAC	TGGTCATCGG	1680
	TGTGGTGGCT	ATCTATATAG	CCGTTAGTCA	TGATTATCCC	AATAATATTC	TGCCCATTTT	1740
35	AGGGCAGCAG	GTCGCCTGGA	TTGCCTTGGG	GCTTGTGATT	GGTTTTGTGG	TCATGCTCTT	1800
	TAATACAGAA	TTTCTTTGGA	AGGTGACCCC	CTTCTATAT	ATTTTAGGCT	TGGGACTTAT	1860
	GATCTTGCCG	ATTGTATTTT	ATAATCCAAG	CTTAGTTGCA	TCAACGGGTG	CCAAAACTG	1920
40	GGTATCAATA	AATGGAATTA	CCCTATTTCA	ACCGTCAGAA	TTTATGAAGA	TATCCTATAT	1980
	CCTCATGTTG	GCTCGTGTCA	TTGTCCAATT	TACAAAGAAA	CATAAGGAAT	GGAGACGCAC	2040
45	GGTCCGCTG	GACTTTTTGT	TAATTTTCTG	GATGATTCTC	TTTACCATTG	CAGTCCTAGT	2100
	TCTTTTAGCA	CTTCAAAGTG	ACTTGGGGAC	GGCTTTGGTT	TTTGTAGCCA	TTTTCTCAGG	2160
	AATCGTTTTA	TTATCAGGGG	TTTCTTGGA	AATTATTATC	CCAGTATTTG	TGACTGCTGT	2220
50	AACAGGAGTT	GCTGGTTTCT	TAGCTATCTT	TATTAGCAAG	GACGGACGAG	CTTTTCTTCA	2280
	CCAGATTGGA	ATGCCGACCT	ACCAAATCAA	TCGGATTTTG	GCTTGGCTCA	ATCCCTTTGA	2340
55	GTTTGCCCAA	ACAACGACTT	ACCAGCAGGC	TCAAGGGCAG	ATTGCCATTG	GGAGTGGTGG	2400
	CTTATTTGGT	CAGGGATTTA	ATGCTTCGAA	TCTGCTTATC	CCAGTTCGAG	AGTCAGATAT	2460
	GATTTTACG	GTTATTGCAG	AAGATTTTGG	CTTTATTGGC	TCTGTCCTGG	TTATTGCCCT	2520
60	CTATCTCATG	TTGATTTACC	GTATGTTGAA	GATTACTCTT	AAATCAAATA	ACCAGTTCTA	2580

-90-

	CACTTATATT TCCACAGGTT TGATTATGAT GTTGCTCTTC CACATCTTTG AGAATATCGG	2640
	TGCTGTGACT GGACTACTTC CTTTGACGGG GATTCCCTTG CCTTTCATTT CGCAAGGGGG	2700
5	ATCAGCGATT ATCAGTAATC TGATTGGTGT TGGTTTGCTT TTATCGATGA GTTACCAGAC	2760
	TAATCTAGCT GAAGAAAAGA GCGGAAAAGT CCCATTCAAA CGGAAAAAGG TTGTATTAAA	2820
10	ACAAATTAAA TAAGGAGAAA ATCATGGTAA AAGTAGCAGT TATGTTAGCT CAGGGCTTTG	2880
	AAGAAATTGA AGCCTTGACA GTTGTAGATG TCTTGCGTCG AGCCAATATC ACATGTGATA	2940
	TGGTTGGTTT TGAAGAGCAA GTAACGGGTT CGCATGCAAT CCAAGTAAGA GCAGATCATG	3000
15	TCTTTGATGG AGATTTATCA GACTATGATA TGATTGTTCT TCCTGGAGGT ATGCCTGGTT	3060
	CTGCACATTT ACGTGATAAT CAGACCTTGA TTCAAGAATT GCAAAGCTTC GAGCAAGAAG	3120
20	GGAAGAACT AGCAGCCATT TGTGCGGCAC CAATTGCCCT CAATCAAGCA GAGATATTGA	3180
	AAAATAAGCG ATACACTTGT TATGACGGCG TTCAAGAGCA AATCCTTGAT GGTCACTACG	3240
	TCAAGGAAAC AGTAGTGGTA GATGGTCAGT TGACAACCAG TCGGGGTCCT TCAACAGCCC	3300
25	TTGCCTTTGC CTACGAGTTG GTGGAGCAAC TAGGAGGGGA CGCAGAGAGT TTACGAACAG	3360
	GAATGCTCTA TCGAGATGTC TTTGGGTAAA AATCAGTAAA ACGGGAGTTA TTCTCTCGTT	3420
30	TTTTATGTGG AAAACTCAGG GAAATCATCG CTTTTTTCAT AAAAAAATGC TATAATGAAG	3480
	GGTATGAAAT ATCACGATTA CATCTGGGAT TTAGGTGGAA CTTTACTGGA TAATTATGAA	3540
	ACTTCAACAG CTGCATTTGT TGAAACATTG GCACTGTATG GTATCACACA AGACCATGAC	3600
35	AGTGTCTATC AAGCTTTAAA GGTCTTCTACT CCTTTTGCGA TTGAGACATT CGCTCCCAAT	3660
	TTAGAGAATT TTTTAGAAAA GTACAAGGAA AATGAAGCCA GAGAGCTTGA ACACCCGATT	3720
40	TTATTTGAAG GAGTTTCTGA CCTATTGGAA GACATTTTAA ATCAAGGTGG CCGTCATTTT	3780
	TTGGTCTCTC ATCGAAATGA TCAGGTTTTG GAAATTTTAG AAAAAACCTC TATAGCAGCT	3840
	TATTTTACAG AAGTGGTGAC TTCTAGCTCA GGCTTTAAGA GAAAGCCAAA TCCCGAATCC	3900
45	ATGCTTTATT TAAGAGAAAA GTATCAGATT AGCTCTGGTC TTGTCATTGG TGATCGGCCG	3960
	ATTGATATCG AAGCAGGTCA AGCTGCAGGA CTTGATACCC ACTTGTTTAC CAGTATCGTG	4020
50	AATTTAAGAC AAGTATTAGA CATATAAGAA AAAGGAATAA GATGACAGAA GAAATCAAAA	4080
	ATCTGCAGGC ACAGGATTAT GATGCCAGTC AAATTCAGT TTTAGAGGGC TTAGAGGCTG	4140
	TTCGTATGCG TCCAGGGATG TACATTGGAT CAACCTCAAA AGAAGGTCTT CACCATCTAG	4200
55	TCTGGGAAAT TGTTGATAAC TCAATTGACG AGGCCTTGCC AGGATTTGCC AGCCATATTC	4260
	AAGTTTTTAT TGAGCCAGAT GATTCGATTA CTGTTGTGGA TGATGGGCGT GGTATCCCAG	4320
60	TCGATATTCA GGAAAAACA GGTGTCCTG CTGTTGAGAC CGTCTTTACA GTCCTTCACG	4380
	CTGGAGGAAA GTTCGGCGGT GGTGGATACA AGGTTTCAGG TGGTCTTCAC GGGGTGGGGT	4440

-91-

CGTCAGTTGT TAATGCCCTT TCCACTCAAT TAGACGTTCA TGTCCATAAA AACGGTAAGA 4500  
5 TTCATTACCA AGAATACCGT CGTGGTCATG TTGTCGAGA TCTTGAAATA GTTGGAGATA 4560  
CGGATAAAAC AGGAACAAC TTTCACTTCA CACCGGACCC AAAATCTTC ACTGAAACAA 4620  
CAATCTTTGA TTTTGATAAA TTAAATAAAC GGATTCAAGA GTTGGCCTTT CTAAATCGCG 4680  
10 GTCTTCAAAT TTCTATCACT GATAAGCGCC AAGGTTTGA ACAAACCAAG CATTATCATT 4740  
ATGAAGGTGG GATTGCTAGT TACGTTGAAT ATATCAACGA GAACAAGGAT GTAATCTTTG 4800  
ATACACCAAT CTATACAGAC GGTGAGATGG ATGATATCAC AGTTGAGGTA GCCATGCAGT 4860  
15 ACACAACGGG TTACCATGAA AAATGTCATG AGTTTCGCCA ATAATATTCA TACACATGAA 4920  
GGTGGAACGC ATGAACAAGG TTTCCGTACA GCCTTGACAC GTGTTATCAA CGATTATGCT 4980  
20 CGTAAGAATA AGTTACTGAA AGACAATGAA GACAATCTAA CAGGGGAAGA TGTTTCGCAA 5040  
GGCTTAACTG CAGTTATCTC AGTTAAACAC CCAAATCCAC AGTTTGAAGG ACAAACGAAG 5100  
ACCAAATTGG GAAATAGCGA AGTGGTCAAG ATTACCAATC GCCTCTTCAG TGAAGCCTTC 5160  
25 TCCGATTTCC TCATGGAAAA TCCACAGATT GCCAAACGTA TCGTAGAAAA AGGAATTTTG 5220  
GCTGCCAAGG CTCGTGTGGC TGCCAAGCGT GCGCGTGAAG TCACACGTAA AAAATCTGGT 5280  
30 TTGGAAATTT CCAACCTTCC AGGGAACTA GCAGACTGTT CTTCTAATAA CCCTGCTGAA 5340  
ACAGAACTCT TCATCGTCGA AGGAGACTCA GCTGGTGGAT CAGCCAAATC TGGTCGTAAC 5400  
CGTGAGTTTC AGGCTATCCT TCCAATTCGC GGTAAGATTT TGAACGTTGA AAAAGCAAGT 5460  
35 ATGGATAAGA TTCTAGCTAA CGAAGAAAT CGTAGTCTTT TCACAGCCAT GGAACAGGA 5520  
TTTGGCGCAG AATTTGATGT TTCGAAAGCC CGTTACCAA AACTCGTTTT GATGACCGAT 5580  
40 GCCGATGTCG ATGGAGCCCA CATTCTGACC CTTCTTTTAA CCTTGATTTA TCGTTATATG 5640  
AAACCAATCC TAGAAGCTGG CTATGTTTAT ATTGCCCAAC CACCAATCTA TGGTGTCAAG 5700  
GTTGGAAGCG AGATTAAAGA ATATATCCAG CCGGGTGCAG ATCAAGAAAT CAAACTCCAA 5760  
45 GAAGCTTTAG CCCGTTATAG TGAAGGTCGT ACCAAACCGA CTATTCAGCG TTATAAGGGG 5820  
CTAGGTGAAA TGGACGATCA TCAGCTGTGG GAAACAACCA TGGATCCCGA ACATCGCTTG 5880  
50 ATGGCTAGAG TTTCTGTAGA TGATGCTGCA GAAGCAGATA AAATCTTTGA TATGTTGATG 5940  
GGGGATCGAG TAGAGCCTCG TCGTGAGTTT ATCGACTCTA GAGGATCCCC GG 5992

(2) INFORMATION FOR SEQ ID NO:40:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 907 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

60

(D) TOPOLOGY: linear



-92-

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TACAAAAGTA GGTGGAGAGG CTGATTATTT GGTCTTTCCA CGAAATCGTT TTGAGTTGGC	60
TCGCGTTGTG AAATTTGCCA ACCAAGAAAA TATCCCTTGG ATGGTTCTTG GCAATGCAAG	120
15 CAATATCATC GTTCGTGATG GTGGGATTCTG TGGATTGTG ATCTTGTTGTG ACAAGCTCAA	180
TAACGTTTCT GTTGATGGCT ATACCATTGA AGCAGAAGCT GGGGCTAACT TGATTGAAAC	240
20 AACTCGCATT GCCCTCCGTC ATAGTTTAAC TGGCTTTGAG TTTGCTTGTG GTATTCCAGG	300
AAGCGTTGGC GGTGCTGTCT TTATGAATGC GGGTGCCTAT GGTGGCGAGA TTGCTCACAT	360
CTTGCACTCT TGTAAGGTCT TGACCAAGGA TGGAGAAATC GAAACCCTGT CTGCTAAAGA	420
25 CTTGGCTTTT GGTTACCGCC ATTCAGCTAT TCAGGAGTCT GGTGCAGTTG TCTTGTCAGT	480
TAAATTTGCC CTAGCTCCAG GAACCCATCA GGTATCAAG CAGGAAATGG ACCGCTTGAC	540
30 GCACCTACGT GAACTCAAGC AACCTTTGGA ATACCCATCT TGTGGCTCGG TCTTTAAGCG	600
TCCAGTCGGG CATTTTGCAAG GTCAGTTAAT TTCAGAAGCT GGCTTGAAAG GCTATCGTAT	660
CGGTGGCGTA GAAGTGTGAG AAAAGCATGC AGGATTTATG ATCAATGTCG CAGATGGAAC	720
35 GGCCAAAGAC TACGAGGACT TGATCCAATC GGTATCGAA AAAGTCAAGG AACACTCAGG	780
TATTACGCTT GAAAGAGAAG TCCGGATCTT GGGTGAAAGC CTATCGGTAG CGAAGATGTA	840
40 TGCAGGTGGT TTTACTCCCT GCAAGAGGTA GTGGGGACCT GACAGAGCCC CGATCGGTTA	900
AGCTATG	907

45 (2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2764 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

55 (iv) ANTI-SENSE: NO

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

-93-

	AGAACCCTTG	GATGCAGCCA	TTCAGAAGAT	TTCTCCAGAA	TTGTTTGACC	AATATGAAAT	60
5	CTTTAAATCA	CGTGAAATGT	TGCTAGAATG	GTCACCAAAG	AATGTTTCATA	AAGCAACAGG	120
	TTTGGCAAAA	CTAATCAGCC	ATCTTGGAAT	CGACCAAAGT	CAAGTGATGG	CTTGTGGTGA	180
	CGAGGCCAAT	GACCTCTCTA	TGATTGAATG	GGCAGGTCTT	GGTGTGCTA	TGCAAAACGC	240
10	TGTTCTTGAA	GTAAAGGCAG	CCGCAAATGT	AGTGACGCCG	ATGACCAACG	ATGAGGAAGC	300
	TGTCGCCTGG	GCTATCGAAG	AATATGTGCT	AAAGGAGAAC	TAAGATATGG	GATTGTTTGA	360
	CCGTCTATTC	GGAAAAAAG	AAGAACCTAA	AATCGAAGAA	GTTGTAAAAG	AAGCTCTGGA	420
15	AAATCTTGAT	TTGTCTGAAG	ATGTTGATCC	TACCTTCACA	GAAGTTGAGG	AAGTTTCTCA	480
	GGAAGAAGCA	GAGGTGAAA	TTGTTGAACA	AGCTGTGTTC	CAAGAAGAGG	AAATCCAAGA	540
20	CACAGTTGAA	GAAAGTCTGG	ATTAGAGCC	AGTTGTAGAA	GTTTCTCAAA	AAGAAGTCGA	600
	AGAATTTCCA	CACTCAGAAG	AAGGGAATAC	TGAGTTTCTA	GAGACTATAG	AAGAAAATAA	660
	TTCTGAAGTT	CTTGAACCAG	AAAGGCCTCA	AGCAGAAGAA	ACCGTTCAGG	AAAAATATGA	720
25	CCGCAGTCTT	AAGAAAACCTC	GTACAGGTTT	CGGTGCCCGC	TTGAATGCCT	TCTTTGCTAA	780
	CTTCCGCTCT	GTTGACGAAG	AATTTTTTCGA	GGAAGTGGAA	GAAGTGTGA	TTATGAGTGA	840
30	TGTTGGTGTC	CAAGTCGCTT	CTAACTTAAC	GGAGGAACTA	CGTTACGAAG	CCAAGCTTGA	900
	AAATGCCAAG	AAACCTGATG	CACTTCGTCTG	TGTCATCATT	GAGAAATTGG	TTGAGCTTTA	960
	TGAAAGGAT	GGTAGCTACG	ATGAAAGCAT	CCACTTCCAA	GATAACTTGA	CAGTTATGCT	1020
35	CTTTGTTGGT	GTGAATGGTG	TTGGGAAAAC	AACTTCTATC	GGAAAGTCTAG	CCCACCGCTA	1080
	CAAACAAGCT	GGTAAGAAGG	TCATGCTGGT	TGCAGCAGAT	ACCTTCCGTG	CGGGTGCAGT	1140
40	AGCTCAGCTA	GCTGAATGGG	GCCGACGAGT	AGATGTTCCA	GTAGTAACTG	GACCTGAAAA	1200
	AGCTGATCCA	GCCAGCGTGG	TCTTTGATGG	TATGGAACGT	GCCGTGGCTG	AAGGTATCGA	1260
	TATTCTCATG	ATTGATACTG	CTGGTCGTCT	GCAAAATAAG	GATAACCTTA	TGGCTGAGTT	1320
45	GGAAAAGATT	GGTCGTATTA	TCAAACGTGT	TGTGCCAGAA	GCACCACATG	AAACCTTCTT	1380
	GGCACTTGAT	GCATCAACAG	GTCAAAATGC	CCTAGTACAG	GCCAAAGAAT	TTTCGAAAAT	1440
50	CACACCTTTA	ACGGGAATTG	TTTGTACTAA	GATTGATGGA	ACTGCTCGAG	GAGGTGTGGT	1500
	TCTAGCCATT	CGTGAAGAAC	TCAATATTCC	TGTAAAATTG	ATTGGTTTTG	GTGAAAAAAT	1560
	CGATGATATT	GGAGAGTTTA	ACTCAGAAAA	CTTTATGAAA	GGTCTCTTGG	AAGGTTTAAT	1620
55	CTAATCAGAA	GCAAAAATCC	TGCAAGGCAT	AACTTGTCAG	GAAATTTTTT	TATTCTAAGC	1680
	GACCATCTTG	ACGATAGGTG	ATATCTGGTT	GCCAAGTCCA	TTTGGCACCG	AATTTTTCAA	1740
60	GTAGGTCAAA	GCTGGCTTGA	GGTCCCATGC	TTCCAGCTTT	ATAGTCATGA	AGTGGGGCAC	1800

-94-

CATTTTCAGC CCAGAGCTTT TCAATACGGT CAATCAACTT CCATGACGCA CAAACTTCAT 1860  
CCCAAGTGGCT AAAGTTAGTT GAGTTGTTAT TTAGGACATC ATAAATCAAT TTTTCGTATG 1920  
5 GTTCTGGAGA AGCACCAGTT GCAGTCGCAT CTGTACGGTA ATCAAGTGAG TTAGGAGCCA 1980  
AGTTAAATTC TTCTCCTACT TGCTTCCCAT TTAGGCTAAG AGAGAAGCCT TCTGTTGGTT 2040  
GAATATAGAT GGTCAAAATA TTTGGAGCAA GTGGTTCTCC AAAGATAGAA TCCATTGTGT 2100  
10 TAAAGACGAT GTTGACATGA GTTCCTTTTT CAGTCAGTCG TTTACCTGTA CGGAAAAAGA 2160  
AAGGAACACC ACGGAATCGA TCGCTGTCTA CAAAGAAGGC ACCAGATGTA AAGGTTTCAG 2220  
15 TTGTTGATTC TGGATTCACA TTTGGCTCGC TACGATAAGA GATGTATTTC ATGCCATCAA 2280  
TCTTACCAGA GCGGTATTGC CCACGGATAA AGTGTTCCTT GAGTTCCTCA TCAGTTGGAT 2340  
GATAGAGGTT TTAAAGACC TTAATCTTTT CAGCACGAAT CTCGTCTTTT GTGAAGCTTG 2400  
20 CTGGTTTGTC CATGGCGAGG AGCGAAAGAA GTTGTAGAGT GTGGTTTGG ACCATGTCAC 2460  
GGAGGACACC GGATTGGTCA TAGTAGCCAC CACGTTCTTC TACACCCAAG CTCCGCAAAG 2520  
25 GTAATTTGAA CATGTGCGAA AAATCCTTGT TCCAAACGTT TTCAAAAATC AAGTTTGCAG 2580  
AGCGAACTGC AAAGATGCTT TGGATCATT CCTTACCAAG ATAATGGTCG ATACGGAAAA 2640  
TTTGTCTTC GTCAAATGTT GCTAGGAGTT CGTCATTCAA CTTGTTTGCA GTTGCCTAAT 2700  
30 CTGTACCAAA TGGTTTTTCA ACGATCAAGC GCTCAAAACC TTTGCCATCG ACTCTAGAGG 2760  
ATCC 2764

35 (2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 3189 base pairs  
    (B) TYPE: nucleic acid  
40      (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

45 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

55 ACACTGTTTA CGATGATGGT CGTATTGATT ACGTGAAAAA CACTTGAGA TCTTGTCTGA 60  
TGCGATTGCA GATGGAGCTA ATGTAAAAGG TTAATTCATT TGGTCATTAA TGGATGTCTT 120  
CTCATGGTCA AACGGTTATG AGAAACGTTA TGGTCTCTTC TACGTAGATT TTGAACTCA 180  
60 AGAACGTTAT CCTAAGAAAT CAGCTCACTG GTACAAGAAA GTAGCGGAAA CTCAGATTAT 240

-95-

	AGACTAGTAG AATTAGTCAT TAGATATAGA ATTTTAGTGA GTCCAAAAGA TGTTCAAAGA	300
	TTTTATCCAA TCTATTTATG AAAAAGTTTA TATTATAAAT TTCGAAAAAT GCTCTCAAAT	360
5	ACCGTGT TTG ACGAGTGAAG AATTGAAAGT CTTGGAAAAT GGTATGTCTC GACTGGTAA	420
	GAATGGATTT GTCATTCAGA TGATGAGCTG GAAGAATTTA AAAATCTATT TTTAAATTTT	480
10	ATCAATCCTG AAGAATGGGA TACTATCTCC TTTGATTCAG ATTTTATGCC GTTTCACAA	540
	TCGTAACCAA TTTCTCAAAA AAGTTAAATC TTATATTTAG TACTCTGTAA AACTCTTATC	600
	TAATCACGTT GCTTATACTC AATGAAAATC ATAGAAAAAA GCATAGTATC AGGTGTTGAA	660
15	ACACCTGATA CTATGCTTTT TATTGTGGGA AGATTACTT TTTTCTTCT GAAATTGAGT	720
	TGTTACCCAG GCTCTTTCAG TTTATTAAGG CTTGATGACT TTAATGTGTT TAGATAGCTT	780
20	AAAAAGGATT GAATCACTTA GTTTAGAATC TGAAACAATA GTATCAAGAT TTGATATATT	840
	ATAAAAAGTA TAAAAATCAA ACTTATTGAA CTTACTATGA TCTGCGAGTA AATATTTTTT	900
	ATTAGAATTA TTAAAGCGA TCGGTTGAGC CTCTCCCTCT TCCTCGCTAA AAGTAGCTAG	960
25	AGCTCCGTTT TGAATACCAT TACAGCTAAC GAAAGCTTTA GAAAATTGGA GATTAGAGAG	1020
	ATTTTGTAGG GTCAATGTGC CAACAAAAGC ACCTGTAATA TCGCGATAAT TTCCACCTAT	1080
30	CAAAATCAAA TCTGTTAATT TTCGTTGCT TAAATCAGA AAAACAGGTA GACTGTTGGT	1140
	TACGACGCGG ATATTGTCAA TAGGCAACTC ACGCGCAAAA AACTCTAATG TTGTTCTGG	1200
	TCCAATGAAA ATAGTTTCTC TTTCTTCTAC TAGACTGCCT GCAAAATGGG CTATTTCTTG	1260
35	TTTTTCTGCC GTTTGGAGGG CTTGTTTTTC AATATTTGAT CGCTCATTAG TCAAAAGGGA	1320
	GTTGGTTCTGA AGTTTTTCAG CTCCACCATG CACACGAATC AGCAAATCTT TATCAGCTAA	1380
40	TTCCTGTAAA TAGCGCCTTG CAGTCATATC TGAAACGGCT ATTTCTGTTA TAATCTGTTT	1440
	AACTGTTATG GTTCCTTTAC TATTTACTAT CTCTAAAATT TTGGCTAATT TTTCTTGTTT	1500
	GAGCATATTA TCACCTCGTT TCCTACTACT ATCTTACCAT AAACAACTC ATCATTCAAA	1560
45	TACAAAAACA ACAAATGAA ACAAAAACAA AAATATCGAA GTTTGTTTTT AAAACTTTTCG	1620
	ATATTTTTGT TGGGTATATA CTTTGATGTT TCTAGTTTAC TTTTGATGA TTGAGAGTGA	1680
50	TGGAGAATTA GTCTAAACCG TAGTTATAGT CATCGTCTTG CATGGCTTCA ACTTCGCCAA	1740
	GAAGGTAACC ATTTCCGACT TGAGAGAAGA AGTCATGGTT GGAAGTTCCT GTTGAAATAC	1800
	CGTTCATAAC GATTGGGTTG ACATCTTCAG CTGAATCTGG GAAAAGTGGG TCTTGTCCTCA	1860
55	TGTTTCATGAG AGCTTTATTG GCATTGTAGC GAAGGAAGGT TTTAACCTCT TCAGTCCAAC	1920
	CAACACCGTC ATAAAGACTC TCTGTGTAGC CTTCTTCATT TTCATAAAGA GTATAGAGTA	1980
60	GGTCGTACAT CCATTCTTTG AGTTTTTCTT GCTCTTCTT AGGCAATTCA TTGAAACCAA	2040
	GTTGGAATTT GTAACCAATG TAGGTCCGT GAACAGACTC GTCACGAATA ATCAATTTAA	2100

-96-

5 TGATTTCTGC AACGTTGGCT AGTTTGTGT TACCGAGATA GTAGAGGGGA GTGAAGAAAC 2160  
CAGAGTAGAA GAGGAAGGTT TCGAGGAAGA CGCTGGCAAC TTTCTTTTCA AGTGGGCTGC 2220  
CGTTTAGGTA GATTTCGTG ACAATCTCAG CCTTCTTTTG TAGGTAAGGA TTGGTATTGG 2280  
TCCATTCGAA AATTTCTTCA ATCTCAGCCT TAGTATTCAA GGTAAGAAAG ATTGATGAGT 2340  
10 AAGATTTAGC GTGGACAGAT TCCATAAATT GGATGTTATT GAAAACAGCT TCCTCATGTG 2400  
GTGTACGGAT GTCTGCGCGA AGGGCTTGAA CCCCAGTTTC AGATTGCATA GTGTCAAGAA 2460  
GGGTAAACC ACCAAAACT TTTCCGACCA AGTCTTCTC TTTGTTAGAT AGCTTTCTCC 2520  
15 AGTCATCCAA GTCGTTTGAT AAGGGAATAC GTGTATCGAG CCAAATTCG TCCGTCAGTT 2580  
TTTCCCAAGT TGATTTGTCG ATGACATCTT CGATGGCATT CCAGTTAATG GCTTTGTAGT 2640  
20 AAGTTTCCAT TTAATCTC TTTCTGTGT TAGTATTGCG AACTCACAAT TATTTCTACT 2700  
TTACCATAAT TCTATAGGAG TATCGCACAA AAAGTCGGAA GCCCGACTTT TAAATGTTA 2760  
CATAAATTAT GTTATGACAT AGTAGATTG ATTTTATCAG TGCTGCTTAG GGAAAAATAA 2820  
25 TGTTTCTATG CTAGAACTA AATCACACAG CTTTCACATT GGTGGCGCC GACTTCTCCA 2880  
CCGTCATCTG TAAAGGTACG GACGTAGTAG ATAGACTTGA TTCCCTTGTT AAAGGCATAG 2940  
30 TTACGAAGGA TGGACAAGTC ACGTGTGCTT TGTTATTTT CCCTCTTCCA TTCGTAAAGG 3000  
CCTTTTGGA TGTCCTACG CATGAAGAGG GTGAGTGAAA GTCCTTGATC CACGTGTTCA 3060  
GTCGCAGCAG CGTAAACATC GATGACTTTA CGCATATCCA TATCGTAGGC AGAAGTGTAG 3120  
35 TAAGGAATGG TTTCTGTAGA CAAGCCAGCA GCAGGGTAGT AGATTTTACC AATTTCTTC 3180  
TCTTGCGT 3189

40 (2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3580 base pairs  
45 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

50 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

60 TTATTGAAGA AGGTGTTAAA GTTGTACAA CAGGAGCAGG AAATCCAAGC AAGTATATGG 60  
AACGTTTCCA TGAAGCTGGG ATAATCGTTA TTCCTGTCGT TCCTAGTGTC GCTTTAGCTA 120

-97-

	AACGCATGGA AAAAATCGGT GCAGACGCTG TTATTGCAGA AGGAATGGAA GCTGGGGGGC	180
5	ATATCGGTAA ATTAACAACC ATGACCTTGG TGCACAGGT AGCCACAGCT GTATCTATTC	240
	CTGTTATTGC TGCAGGAGGA ATTGCGGATG GTGAAGGTGC TGC GGCTGGC TTTATGCTAG	300
	GTGCAGAGGC TGTACAGGTG GGGACACGGT TTGTAGTTGC AAAAGAGTCG AATGCCCATC	360
10	CAAAC TACAA GGAGAAAATT TTAAGCAA GGGATATTGA CACTACGATT TCAGCTCAGC	420
	ACTTTGGTCA TGCTGTTCGT GCTATTAAAA ATCAGTTGAC TAGAGATTTT GAACTGGCTG	480
15	AAAAAGATGC CTTTAAGCAG GAAGATCCTG ATTTAGAAAT CTTTGAACAA ATGGGAGCAG	540
	GTGCCCTAGC CAAAGCAGTT GTTCACGGTG ATGTGGAGGG TGGCTCTGTC ATGGCAGGTC	600
	AAATCGCAGG GCTTGTCTTCT AAAGAAGAAA CAGCTGAAGA AATCCTAAAA GATTTGTATT	660
20	ACGGAGCCGC TAAGAAAATT CAAGAAGAAG CCTCTCGCTG GACAGGAGTT GTAAGAAATG	720
	ACTAAACAG CCTTTTATT TGCTGGTCAA GGTGCCCAGT ATCTAGGGAT GGGACGGGAT	780
25	TTCTATGATC AGTATCCGAT TGTTAAAGAA AC GATTGATC GAGCGAGTCA GGTGCTAGGT	840
	TATGATTTGC GTTATCTCAT CGATACGGAA GAAGACAAAC TCAATCAGAC CCGCTATACG	900
	CAACCAGCCA TTCTAGCGAC TTCGGTTGCT ATCTACCGTT TATTGCAAGA AAAGGGCTAT	960
30	CAGCCTGATA TGGTTGCTGG TTTGTCTCTT GGAGAATACT CTGCCTTGGT GGCAAGCGGC	1020
	GCCTTGGATT TTGAAGATGC GGTTCCTTG GTAGCTAAGC GTGGAGCCTA TATGGAAGAA	1080
35	GCGGCTCTG CTGACTCTGG CAAGATGGTA GCAGTTCTCA ATACGCCAGT AGAGGTCATT	1140
	GAAGAAGCCT GTCAAAAAGC TTCTGAACTT GGAGTGGTTA CTCCAGCCAA CTATAACACA	1200
	CCTGCACAAA TCGTCATTGC TGGAGAAGTG GTTGCAAGTG ATCGAGCGGT TGAAC TTTTG	1260
40	CAAGAAGCAG GTGCCAAACG CTTGATTCTT CTTAAGGTGT CAGGTCCCTT TCACACCTCT	1320
	CTCCTTGAAC CTGCTAGCCA GAACTAGCT GAACTCTGG CTCAGGTAAG TTTTTCAGAT	1380
45	TTTACTTGTC CCCTAGTCGG CAATACAGAA GCTGCTGTGA TGCAAAAAGA GGACATTGCT	1440
	CAGCTCTTGA CGCGTCAGGT CAAGGAACCC GTTCGTTTCT ATGAAAGTAT TGGGGTCATG	1500
	CAAGAAGCAG GCATAAGCAA CTTTATCGAG ATTGGACCGG GGAAAGTCTT GTCAGGTTTT	1560
50	GTAAAAAAA TTGATCAAAC TGCTCACTTA GCTCATGTGG AAGATCAAGC GAGTTTAGTA	1620
	GCACTTTTAG AAAAATAGAC TAAAATAAGT AGAAGTTTGT AAAGGAAAAA AATGAAACTA	1680
55	GAACATAAAA ATATCTTTAT TACAGGTTCTG AGTCGTGGAA TTGGTCTTGC CATCGCCAC	1740
	AAGTTTGCTC AAGCAGGAGC CAACATTGTC TTAAACAGTC GTGGGGCAAT CTCAGAAGAA	1800
	TTGCTCGCTG AGTTTTCAAA CTATGGTATC AAGGTGGTTC CCATTTCAGG AGATGTATCA	1860
60	GATTTTGCAG ACGCTAAGCG TATGATTGAT CAAGCTATTG CAGAACTGGG TTCAGTAGAT	1920

-98-

	GT TTTGGTCA ACAATGCAGG GATTACCCAA GATACTCTTA TGCTCAAGAT GACAGAAGCA	1980
	GATTTTGAAA AAGTGCTCAA GGTCAATCTG ACTGGTGCCT TTAATATGAC ACAATCAGTC	2040
5	TTGAAACCGA TGATGAAAGC CAGAGAAGGT GCTATCATT AATATGTCTAG TGTTGTTGGT	2100
	TTGATGGGGA ATATTGGTCA AGCTAACTAT GCTGCTTCTA AGGCTGGCTT GATTGGCTTT	2160
	ACCAAGTCTG TGGCACGCGA GGTGCTAGT CGGAATATAC GAGTCAATGT GATTGCTCCA	2220
10	GGAATGATTG AGTCTGATAT GACAGCTATC TTATCAGATA AGATTAAGGA AGCTACACTA	2280
	GCTCAGATTC CGATGAAAGA ATTTGGGCAG GCAGAGCAGG TTGCAGATT GACAGTATT	2340
15	TTAGCAGGCC AAGATTATCT AACTGGTCAA GTGATTGCCA TTGATGGTGG CTTAAGTATG	2400
	TAGCGAAAGC TAGAGGTGAA AAGAATGAAA CTAAATCGAG TAGTGGTAAC AGGTTATGGA	2460
	GTAACATCTC CAATCGGAAA TACACCAGAA GAATTTTGG AATAGTTTAGC AACTGGGAAA	2520
20	ATCGGCATTG GTGGCATTAC AAAATTGAT CATAGTGA CTGATGTGCA TAATGCGGCA	2580
	GAAATCCAAG ATTTTCCGTT CGATAAATAC TTTGTAAAA AAGATACCA CCGTTTGTG	2640
25	AACTATTCTT TATATGCCTT GTATGCAGC CAAGAGGCTG TAAACCAGCC AATCTTGATG	2700
	TAGAGGCTCT TAATAGGGAT CGTTTGGTG TTATCGTTGC ATCTGGTATT GGTGGAATCA	2760
	AGGAAATTGA AGATCAGGTA CTTGCTCTT ATGAAAAAG ACCCAAACGT GTCAAACCAA	2820
30	TGACTCTTCC AAAAGCTTTA CCAAATATGG CTTCTGGGA TGTAGCCATG CGTTTGGTG	2880
	CAAACGGTGT TTGTAAATCT ATCAATACTG CCTGCTCTT ATCAAATGAT GCGATTGGGG	2940
35	ATGCCTTCCG CTCCATTAAG TTTGGTTTCC AAGATGTGAT GTTGGTGGGA GGAACAGAAG	3000
	CTTCTATCAC ACCTTTTGCC ATCGCTGGT TCCAAGCCTT AACAGCTCTC TCTACTACAG	3060
	AGGATCCAAC TCGTGCTTCG ATCCCATTTG ATAAGGATCG CAATGGGTTT GTTATGGGTG	3120
40	AAGGTTCAAG GATGTTGGT CTAGAAAGTC TTGAACACGC TGAAAAACGT GGAGCTACTA	3180
	TCCTGGCTGA AGTGGTTGGT TACGGAAATA CTTGTGATGC CTACCACATG ACTTCTCCAC	3240
45	ATCCAGAAGG TCAGGGAGCT ATCAAGGCCA TCAACTAGC CTTGGAAGAA GCTGAGATT	3300
	CTCCAGAGCA AGTAGCTATG TTAATGCTCA CGGAACGTCA ACTCCTGCCA ATGAAAAAGG	3360
	AGAAAGTGGT GCTATCGTAG CTGTTCTTGG TAAGGAAGTA CCTGTATCAT CAACCAAGTC	3420
50	TTTACAGGA CATTTGCTGG GGGCTGCGGG TGCAGTAGAG CTATCGCACC ATCGAGCTAT	3480
	GCGTCATACT TTGTACCATG CCAGCTGGGC AAGTGAGGTA TCAGATATAT CGAGCTAATG	3540
55	TCGTTATGGC AGGTTTGAGA AGAATTCATA CGTATTCAAA	3580

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1780 base pairs

(B) TYPE: nucleic acid

60

-99-

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

15	ATGTCCGCAA GAATGTGATT AATCAGCAAT CATCCTTGAT CGGAGATGAA TCGATCTTGG	60
	CTTTTGGAGT GAACCAGCCT TTAGCGGAT TTGGTGTAA AGGAGAAAGA CAGCAACAGC	120
	ATCAGCCTAT GACCTCTATG TACTGAAGCG ACCACTTCCC CAAGTAGGAC CTCGATGTCA	180
20	TTTTAGATAG TCAAAATCAG GCTGTCTGCA TTGTCGAAAT TACAAAGGTT TCTGTTGAAC	240
	TCTTCAATCA AGTTTCTGCG CAACATGCCT TTAAGGAAGG TGAGGGAGAC AAATCACTTG	300
25	CCTATTGGCG CCAGGTTTAT GAGGACTTTT TCACAGACTG TTTGGGTGAA GTAGGGCTGA	360
	CTTTTACACC TGAAAGCAAG GTTGTTTTAG AAGAATTCG CAAGGTCTAC CCACTGTAGA	420
	CTATTAGAAG GAAGAAAGTT TTGGAAATCG CTGTCCAATC CTTTTTCTC AAGCAAATA	480
30	TGATATAATA AGTTTGTGTTG AAGAAGAGCA GCAGCTCTTA AACTTAGAAT AGGAGAAAAC	540
	TATGCAAGCA GTTGAACATT TTATTAAGCA ATTTGTTTCT GAACATTATG ATTTATTTTT	600
35	AGATTTGAGT CGTGAGACCA AGACTTTTTT TGGGAAAGTG ACCATCACTG GTCAAGCACA	660
	GAGTGACCGC ATCTCCCTCC ACCAAAAAGA CTTGGAATC ACCTCTGTAG AAGTTGCAGG	720
	TCAAGCTCGT CCATTACAG TTGACCATGA CAATGAAGCC CTTATATCG AATTGGCTGA	780
40	GGCTGGTCAA GTTGAATTGG TTCTTGCTTT TTCTGGTAAA ATTACAGACA ACATGACAGG	840
	GATTTACCTT TCTTATTATA CAGTTGATGG AGTCAAGAAG GAGGTCTTGT CTAATCAGTT	900
45	CGAGAGCCAT TTTGCGCGCG AAGCTTTCCC ATGTGTGGAT GAGCCTGAAG CCAAAGCAAC	960
	TTTTGACCTC TCTCTTCGCT TTGACCAAGC AGAAGGTGAA TTGGCCTTGT CAAACATGCC	1020
	AGAAATCGAT GTTGAAAACC GTAAGGAAAC AGGTATCTGG AAGTTTGAGA CAACACCTCG	1080
50	CATGCTTCTT TACTTGTTGG CCTTGTTGTC TGGTGATTTG CAAGGGGTGA CCGCTAAAC	1140
	TAAAAATGGT ACCCTGGTAG GTGTCTACTC AACCAAAGCA CATCCACTTT CAAATCTTGA	1200
55	TTTCTCACTG GATATCGCTG TTCGCTCTAT CGAGTTTAC GAAGATTACT ATGGAGTTAA	1260
	GTACCCAATT CCTCAATCTC TCCACATCGC CCTTCTGAC TTCTCAGCTG GTGCTATGGA	1320
	AAACTGGGGT CTTGTGACCT ACCGTGAAGT TTAATTGGTT GTCGATGAGA ACTCTACATT	1380
60	TGCTAGCCGT CAACAAGTTG CCCTTGTTGT GGCCCATGAA TTGGCTCACC AATGGTTTGG	1440



-100-

GAACCTCGTG ACTATGAAAT GGTGGGATGA CCTTTGGCTC AATGAAAGTT TCGCTAATAT 1500  
GATGGAATAC GTCTGTGTGG ATACCATCGA ACCAAGCTGG AATATCTTTG AAGATTTCCA 1560  
5 AACAGGTGGA GTACCTCTTG CTCTTGAACG TGACGCTACT GATGGCGTTC AGTCTGTCCA 1620  
CGTCGAAGTT AACATCCAG ATGAGATCAA TACACTCTTT GACGGCGCTA TCGTCTATGC 1680  
10 AAGGAAGCGT CTCATGCACA TGCTTCGCGT TGCTAGAGAT GCTGATTGT AAGGTTGCAC 1740  
GCCTACTTTG GAAACACCAT ACAGCACACC ATTGGAGTGA 1780

## (2) INFORMATION FOR SEQ ID NO:45:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 671 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
20 (D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
25 (iv) ANTI-SENSE: NO

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GTCTTTTGTA GCGAGGCCAG TGTCTTTTGC CCATCATTTG TCAGGCAGAT AAAACTAGAG 60  
35 CGTCTATCTT GATGGCAACA CATGCGACTG AGTAGACCGC AATTTTtagc TTCCAAGCGA 120  
GCCACCATCC TAGAAACTGC GCTCGGGCTC AGATGAAGCT TATCTGGCAG GTCAATCTGG 180  
CGTAGAGATT TTTCTTCAGC CAAGTCCAGA TAGTAGAGCA GGTAGAACTC TTTCAAGGTC 240  
40 AGACTTTGCT CGCTCTGTTG GGCAATGGTC TCTTCCAAGA GACTTTCAAT TTCTTTCTGA 300  
CGCCGATTGA AGTCAAACCA TTTTTCAAA TAGGTCATAG TGTCTCCTTT CTTTTTAGAG 360  
45 TCATAAAATA GAAGAAAGTC CATTACGGG CAGTCTCTGC GTCACAAGAT GATTGCGCAT 420  
GCAATAATTA TACTACTTTT CAAGAATGCT GGCAAGCTCT GTTTTTTAGT GGTTTTCTTT 480  
TTTACTGTCT ATATTTTTTG TAAAATCAA CTTTTACTTG GATGAAGGTT TTGGCTTCAC 540  
50 GTAGGAGTTG AAGAAGGGTG GCGCGGGTTT CAATTCTTCT CTTGTCTTGG GCAGACTGCG 600  
GTTCGGAAG ACTTCCAGAT AACGTTCAAT TTCATCTAGC AATCAGAGCA GGATTGGTCT 660  
55 GGCTCAGTGA C 671

## (2) INFORMATION FOR SEQ ID NO:46:

60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1557 base pairs  
(B) TYPE: nucleic acid

-101-

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

15	TTTCAGCTCA CAAATATAGG TCGGATGAGC CACTTCCTTA CGAACACGCG CATCAAAAGC	60
	ATCTAGCTCC TCACGTGAAA AAGCATCCTG CAAACTATAA AGAGGATACT GATGACTGTA	120
	TTTTTCAAAA CCATCTAAAA CCTTGCCACC AACACGATGA GTCGGACTGT CTGCTAGCAC	180
20	TTGCTCTGGA TAAGCAGTTT CTAACGAC CACTCACGG TAAAGGCGGT CATACTCACT	240
	GTCTGAAACC GAGGGATTAT CGCTGGTATA GTACTCAGTC GCATAGCGAT TGAGCAAAGC	300
25	GACTAACTCA TTCATTCTTT TATTCATAAG ACCATTTTAC CATAAAACAA GCCCTCCTCA	360
	CAAACGAGAA GGGCGGAAAA AACACTTAGT TTGAAATTAT TTTTGAAACT CAAGCAACCT	420
	TATATCAATT TTTCAAAATG AGTTCGAACA TAAATAAAG ATATACAAGA CAAGATGATA	480
30	ACACCACTTC CAATTATCAG GAAAGAAGAG AGATGTACAC TTGGCAAGAC TGTCAATAAT	540
	CCTTTTGCAA TAGGCATAAA TAGAATAGCT AAGGTAAAAA TTGTACTCAG TACTCTTCCA	600
35	AGAAATTCGC TCTCAACCTT GGTTCGTACT TGAGTAAAAA AGTGAATATT AAAAATCGTC	660
	ATAACAATT CACAACTAA ATTTCCAGAA AAGGAAAGAA AAGTTGGAAG TGGTAATCCC	720
	ATCATAAAAA CTCCGACACC TGTCAAAGCC AGTAAATCA AAAGATTATA AATATTAGCT	780
40	TTAATTTTAC TAGCTAGAAG AGCCCCAATG ATGGAACCAA TAGCCCCCAT AGTTAAAATA	840
	CTTGATAGG CTCCTTCTGA CCCGTAAAGC TGATTCGAAA AGGGAAGTAG AAATTCAAAA	900
45	GCTGCAAAAA AGAAATTAAC GCTGGAAGCT ACCAGCAAAA GGAAGAAAAT TTCTTGCTGA	960
	TGCCAGATAT AGTGTAACCC ATCCTTGATA TCTACAAAAA TATCTCTCCC AGTAAAAGCC	1020
	TTTTTCTCTT GAACTTTTGC TTCCTCTTTT GGAAGGAAAG CCACTAGAAC AAAAGCAATG	1080
50	AAAAAAGTCA GCGAGTCTAG CAGTAGCGTC ATATGGAGAC TTGCAAACCTG TAAAACAAGG	1140
	AAGGAAAGAA CAGGAGAGCT AACACCTACA ACCTGCAAAA CCAGCTCTAA GCGAGAATTA	1200
55	TAGATCACAA TCTCATTTTT CTCCACCACT TCAGTTATGA TAGCTTTATT GGCTGTGCGA	1260
	GAAAAGGCAA AAGCAATAGC CTGCACAATG TTAGCAACAA TCAAAGCGCC AATCATCCAG	1320
	CTATCATTCCT TTATGAAAGA AATAGCCAGA CAAAGAATCC CACAAACAAG ATCTGCCGTC	1380
60	ATTAAAATCT TACGACGAGA AAAACGGTCT GAAATAACTC CGCCAAAGGG ATTGACGAGA	1440

-102-

ATAGATGTGA CGAGCTCAGA AATCTGATAC ATTCCTAAAA CTGTCTGTCC TATAGTCCCC 1500

ATAGAAGCCA ACCAGACACT ATTTCCATAA TCATAGAGCA TATTCCCATT TTATTGA 1557

5

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 658 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

25 CTTATTTGGT TTGGGAATTC GTCATGTCGG AAGCAAGGCT AGTCAGCTTT TACTTCAATA 60

TTTCCATTCA ATTGAAAATC TGTATCAGGC AGATTCAGAG GAAGTGGCTA GTATTGAAAG 120

30 TCTAGGTGGC GTGATTGCCA AAAGTCTTCA GACTTATTTT GCGGCAGAAG GCTCTGAAAT 180

TCTGCTCAGA GAATTGAAAG AACTGGGGT CAATCTGGAC TATAAAGGAC AGACGGTAGT 240

AGCGGATGCG GCCTTGTCAG GTTTGACCGT GGTATTGACA GGAAAATTGG AACGACTCAA 300

35 GCGCTCAGAA GCTAAAAGTA AACTCGAAAG TCTGGGTGCC AAAGTGACAG GTAGTGTTC 360

TAAAAGACC GACCTCGTCG TGGTAGGTGC AGACGCTGGA AGTAACTGC AAAAAGCACA 420

40 AGAACTTGGT ATCCAGGTCA GAGATGAGGC ATGGCTAGAA AGTTTGTAAT GGATCGTTTA 480

AAAACAGAGT TTAGAGAATA TGACTATGTC TGTTAATTGA GACGAGATTG ACAAAAATTT 540

ATTAGTGAAA TAGGAAACAA AGTAAAAGG AAAAATAAAA AATGTATACT ACCCTATGCG 600

45 CATTCAATTAC CATCGTAAGA ATGGAGAATA TGACCTTGCT CCTTTGTAAA AGTCAGGA 658

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 2474 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

55

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

60

-103-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

5	ACAATCGATC AGACAGTCAA TCGATTCTA AAATGTTAG AGTAGAGATG TACCTATTCT	60
	AGTTCAATAT ACTATATAAC TGAAAATTTA GATAAATTAG TTTTGGAAT GACTAACCAA	120
	AGATATCCAA AGTAGTCTAA AATTGTCTAT ACTTTATGAG TGTTTtagTT AGGAAAAAGG	180
10	CTTGTTGTCT ATAATTGGCG CATTAGTCTA GATTTTATTT ATAGAAAATG TTATAATAGA	240
	CTGTATTTAA AAAATTTTAA GGAGAAATGA CAGAATGTCT GTATCATTG AAAACAAAGA	300
15	AACAAACCGT GGTGTCTTGA CTTTCACTAT CTCTCAAGAC CAAATCAAAC CAGAATTGGA	360
	CCGTGTCTTC AAGTCAGTGA AGAAATCTCT TAATGTTCCA GGTTCGGTA AAGGTCACCT	420
	TCCACGCCCT ATCTTCGACC AAAAATTGG TGAAGAAGCT CTTTATCAAG ATGCAATGAA	480
20	CGCACTTTTG CCAAACGCTT ATGAAGCAGC TGTAAGAAGAA GCTGGTCTTG AAGTGTTGC	540
	CCAACCAAAA ATTGACGTAA CTTCAATGGA AAAAGGTCAA GACTGGGTTA TCACTGCTGA	600
25	AGTCGTTACA AAACCTGAAG TAAAATTGGG TGAAGTACAA AACCTGAAG TATCAGTTGA	660
	TGTAGAAAA GAAGTAACTG ACGCTGATGT CGAAGAGCGT ATCGAACGCG AACGCAACAA	720
	CCTGGCTGAA TTGGTTATCA AGGAAGCTGC TGCTGAAAAC GCGGACACTG TTGTGATCGA	780
30	CTTCGTTGGT TCTATCGACG GTGTTGAATT TGACGGTGGA AAAGGTGAAA ACTTCTCACT	840
	TGGAATTGGT TCAGGTCAAT TCATCCCTGG TTTCGAAGAC CAATTGGTAG GTCCTCAGC	900
35	TGGCGAAACC GTTGATGTTA TCGTAACATT CCCAGAAGAC TACCAAGCAG AAGACCTTGC	960
	AGGTAAAGAA GCTAAATTCG TGACAACCTAT CCACGAAGTA AAAGCTAAAG AAGTTCAGC	1020
	TCTTGACGAT GAACTTGCAA AAGACATTGA TGAAGAAGTT GAAACACTTG CTGACTTGAA	1080
40	AGAAAAATAC CGCAAAGAAT TGGCTGCTGC TAAAGAAGAA ACTTACAAAG ATGCAGTTGA	1140
	AGGTGCAGCA ATTGATACAG CTGTAGAAAA CGCTGAAATC GTAGAACTTC CAGAAGAAAT	1200
45	GATCCATGAA GAAGTTCACC GTTCAGTAAA TGAATTCCTT GGGAACTTGC AACGTCAAGG	1260
	GATCAACCCT GACATGTACT TCCAAATCAC TGGAATACT CAAGAAGACC TTCACAACCA	1320
	ATACCAAGCA GAAGCTGAGT CACGTACTAA GACTAACCTT GTTATCGAAG CAGTTGCCAA	1380
50	AGCTGAAGGA TTTGATGCTT CAGAAGAAGA AATACAAAA GAAGTTGAGC AATTGGCAGC	1440
	AGACTACAAC ATGGAAGTTG CACAAGTTCA AAATTGCTT TCAGCTGACA TGTGAAACA	1500
55	TGATATCACT ATCAAAAAAG CTGTTGAATT GATCACAAGC ACAGCAACAG TAAAAAATC	1560
	TTAATAAACA GAAAACCCAC CTGAATTGGT GGGTTTCTG ATGCACTATT TTCCAAAAAT	1620
	CTCTTTGAGG TCTGTGTCTG TAATCCCAAT CATGGCTGGG ATGCGGTCCC AGTTTCTTC	1680
60	GGTTAGGATG TAGGATTGTT CAGAGGCACT TGATGTGACT GTTTCAGAGA CAGCTTGTG	1740

-104-

CTTTTCTTCA ACATTCTCCA GTAGATCACT GAAGCGTTCA ATCAGATAGG TTTTTCGGGC 1800  
AGTTCCGATG TGTGCGGTAG CATAGTCGAA GGCTTGTAAT TCGCCTAGTA AGATGAGTTT 1860  
5 GCTTTTGGCA CGTGTAATGG CTGTGTAGAT GAGATTTTCGC TCCAGCATAC GTCGGCTAGC 1920  
ACTAGTAATC GGTAGGATGA CAACTGGGAA CTCACTTCCC TGAGACTTAT GAATACTCAT 1980  
10 GGCATAGGCC AAGCGAATCT TGTACCATTG GTTACGGGGG TAAGAGACTT CATTACTATC 2040  
AAAATCAATG ACAATCTCGT CTTGTTTCGA TTCGGTGTAT TTACCAGGAA TCAGGTCTGT 2100  
GATAGCTCCT AAATCCCCAT TAAAGACATT GATTTTCAGCA TCGTTAACCA AATGAATGAC 2160  
15 CTTGTCTCTC TTACGATAGT GACACTGAGG AGCTTCAAAA CTGAGTTGAT CTTTTTGTGG 2220  
GGGATTGAGC AGGTCTTGCA TGAGCTGATT GATAGCATCA ATCCCTGCCG TCCCTCGGTA 2280  
20 CATAGGAGCC AGAACTTGA TATCACGGGC GGGGAATACCA TTTCTGAGGG CGGCACCTAA 2340  
GATTTTTTCA ATGGTGGCAG GAATATGGCC ACTAGCAATT TCAAAGTAGG AACGGTCAGC 2400  
TTTTTTTTTG GTGAAATCAG CTGGCAAGAT GCCCTGTCGA ATCTGACTAG CTAGGGTGAC 2460  
25 GATGGTTGAT TCTT 2474

## (2) INFORMATION FOR SEQ ID NO:49:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 716 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

40 (iv) ANTI-SENSE: NO

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

ATCAAAATTA ACGTATTCTT TTTGAAGTTC AAGAACTTCT TCCATTGTTG AGCATTCTGT 60  
AAGGGCACGG TTTGCGTACT CTTCCATCTT AGCTGTGTCG AGTTTCTTCA TCAAGCTGCG 120  
50 TGTACGAAGT ACAGATGTTG CTGACATAGA GAACTCATCC AAGCCCATTC CGACAAGAAG 180  
TGGAACAGCT TGTGTTGTCAC CAGCCATCTC ACCACACATA CCAGCCCATT TACCTTCAGC 240  
55 GTGAGCTGCT TTGATCACAT TGTTAATCAA GCGTAGGATT GATGGGTTGT ATGGTTGGTA 300  
AAGGTATGAA ACTTGTTTCGT TCATACGGTC TGCTGCCATT GTATATTGGA TCAAGTCGTT 360  
TGTACCAATT GAGAAGAAGT CAACTTCTTT AGCAAATTGG TCTGCAAGCA TAGCCGCTGC 420  
60 AGGAATCTCG ATCATGATAC CAACTTGAAT GTTATCCGCA ACTGCAACAC CTTAGCAAG 480

-105-

AAGGTTTGCT TTTTCTTCAT CAAAGACTGC TTTCGCTGCA CGGAATTCTT TCAAGAGCGC 540  
AACCATTGGG AACATGATAC GCAATTGACC GTGAACAGAC GCACGAAGAA GAGCACGGAT 600  
5 TTGTGTGCGG AACATAGCAT CTCCAGTCTC AGAGATAGAG ATACGAAGAG CACGGAATCC 660  
AAGGAATGGG TCATTCTGA GGCATATCGA AGTAAGGAAG TCCTTATCTC CACCGA 716

10 (2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 962 base pairs  
(B) TYPE: nucleic acid  
15 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

30 AGTAACCTAA ATCAATTATG GTGTTATGAG TCTTGGTGTG CCCAAAGTGC TGACGTA ACT 60  
ATCTCAGCTG AAGGTGCAGA TGCAGATGGC CTATCGCTGC AATCTCAGAA ACAATGGAAA 120  
AAGAAGGATT GGCATAAGGG AAATGACAGA AATGCTTAAA GGAATCGCAG CATCTGACGG 180  
35 TGTTGCAGTT GCAAAGCAT ATCTACTCGT TCAGCCGGAT TTGTCATTG AGACTATTAC 240  
AGTCGAAGAT ACAAACGCAG AAGAAGCTCG CCTTGATGCC GCTCTACAGG CATCACAAGA 300  
CGAGCTTTCT GTTATTCGCG AGAAAGCAGT AGGTACGCTC GGTGAAGAAG CAGCTCAAGT 360  
40 TTTTGATGCT CACTTAATGG TTCTTGCTGA CCCAGAAATG ATCAGCCAAA TCAAGGAAAC 420  
TATCCGTGCG AAGAAAGTGA ATGCAGAAGC AGGTCTGAAA GAAGTTACAG ATATGTTTAT 480  
45 CACTATCTTT GAAGGCATGG AAGACAACCC ATACATGCAA GAACGCGCAC GGATATCCGC 540  
GACGTGACAA AACGTGTATT GGCAAACCTT CTTGGTAAAA AATTGCCAAA CCCAGCTTCT 600  
ATCAATGAAG AAGTGATTGT GATTGCGCAT GACTTGACTC CTTAGATAC AGCTCAATTG 660  
50 GACAAAAACT TTGTAAAAGC TTTTGTAACC AACATTGGTG GACGTACAAG CCACTCAGCT 720  
ATCATGGCAC GTACACTTGA AATTGCTGCT GTATTAGGTA CAAACAACAT CACTGAAATC 780  
55 GTTAAAGACG GTGACATCCT TGCTGTTAAC GGGATCACTG GAGAAGTGAT TATCAACCCA 840  
ACAGATGAAC AAGCGGCAGA ATTTAAAGCA GCTGGTGAAG CCTATGCGAA CAAAAGCTG 900  
AATGGGCACT TTTGAAAGAT GCTCAACAGT GACTGCTGAC GGTAACACTC GAGTTGGCTG 960  
60 CC 962

-106-

## (2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:  
5 (A) LENGTH: 2702 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

20 GATCGTTTCC GTGGCTTGAT CGGAAGCATG TTTGACGAAT AAAGAGGAAA AATAAATTAT 60  
GACATTTTCA TTTGATACAG CTGCTGCTCA AGGGGCAGTG ATTAAAGTAA TTGGTGTCGG 120  
25 TGGAGGTGGT GGCAATGCCA TCAACCGTAT GGTGACGAA GGTGTTACAG GCGTAGAATT 180  
TATCGCAGCA AACACAGATG TACAAGCATT GAGTAGTACA AAAGCTGAGA CTGTTATTCA 240  
GTTGGGACCT AAATTGACTC GTGGTTTGGG TGCAGGAGGT CAACCTGAGG TTGGTCGTAA 300  
30 AGCCGCTGAA GAAAGCGAAG AAACACTGAC GGAAGCTATT AGTGGTGCCG ATATGGTCTT 360  
CATCACTGCT GGTATGGGAG GAGGCTCTGG AACTGGAGCT GCTCCTGTTA TTGCTCGTAT 420  
35 CGCCAAAGAT TTAGGTGCGC TTACAGTTGG TGTGTAAACA CGTCCCTTTG GTTTTGAAGG 480  
AAGTAAGCGT GGACAATTTG CTGTAGAAGG AATCAATCAA CTTGCTGAGC ATGTAGACAC 540  
TCTATTGATT ATCTCAAACA ACAATTTGCT TGAAATTGTT GATAAGAAAA CACCGCTTTT 600  
40 GGAGGCTCTT AGCGAAGCGG ATAACGTTCT TCGTCAAGGT GTTCAAGGGA TTACCGATTT 660  
GATTACCAAT CCAGGATTGA TTAACCTTGA CTTTGCCGAT GTGAAAACGG TAATGGCAAA 720  
45 CAAAGGGAAT GCTCTTATGG GTATTGGTAT CGGTAGTGGA GAAGAACGTG TGGTAGAAGC 780  
GGCACGTAAG GCAATCTATT CACCACTTCT TGAAACAACCT ATTGACGGTG CTGAGGATGT 840  
TATCGTCAAC GTTACTGGTG GTCTTGACTT AACCTTGATT GAGGCAGAAG AGGCTTCACA 900  
50 AATTGTGAAC CAGGCAGCAG GTCAAGGAGT GAACATCTGG CTCGGTACTT CAATTGATGA 960  
AAGTATGCGT GATGAAATTC GTGTAACAGT TGTCGCAACG GGTGTTTCGTC AAGACCGCGT 1020  
55 AGAAAAGGTT GTGGCTCCAC AAGCTAGATC ACCGCGCCTA GGATAACAAT TTTAGCAATC 1080  
AAGATAAACC AAAACATCAT AACAACAAGA AGAACGGAAC CTAAAATTCG GACATCCACC 1140  
AAATGATGGA CATAGTAATT GAGATAACTA GAGAACAGAG TTAGTACACC TAAAATCACC 1200  
60 AAGAGAACAA AGGCACTGCC TGGTAGGGTA TAGCTAATTT TCCTGTTAGA TAGATTGGGA 1260

-107-

AGAAAATAAT AAAGCATGAC CAAGATAGCA AAGAGGAGGG CGTAAATCAG AGGACCTGCC 1320  
AACCCTTGTA AAGCCTGATA GATAATGCCA TCTTTTGTCC AATAATGAGC AAGTAAAGCC 1380  
5 AAAATCATCT GACCAAATAA GATCAAAAAC AAGGCAAACG CAAAGAGGAA CTGCAAGCCA 1440  
AAACTGACTA GGAGACTTAG CATCTGATGG GAAATAAGTC CACGACTCTT TTCGACGCCA 1500  
10 TAAGCCTTGT TAAAAGCTTT TTGCAAGAAA TTTATAGATT TTGAAAACT CCATAACGCC 1560  
GATAAACAG AAAAAGCTCA TAAACCTGTT GAAGGTTGCG TCAAAGACTT CTCTGGCTAT 1620  
TTTTTCCACA CCTTCATAGA GGCTTGGGGG CAGGACGTCT TTCATAAAGC CCAGAAATTC 1680  
15 TCCCACAGGA ATCTGAAAAT AGGGGAGGAT ATTGACCACC ACCAAAAGCA GGGGGAAAAT 1740  
CGAAATCAAC CAATAGTACG CTAAGTACG ACTGGTCAAA CTCACTATCT GATGCTTGAT 1800  
20 AATAATGCAA AAAAGCTTTT AATAAAGGCT TGTCTATCAG CTCTTTCCAC CACTTTTTTCA 1860  
TGTCATACTC CTTCAATTAT AATCTTATAC TCAATGAAA TCAAAGAGCA AACTAGAAAG 1920  
CTAGCCGCAA GCTGCTCAAA AACTGTTTTT GAGGTTGTAG ATAAGACTGA CGAAGTCAGT 1980  
25 CACATACATA CGGTAAGGCG ACGCTGACGT GGTGTTGAAGA GATTTTCGAA GAGTATTAAC 2040  
TAATTTCTTC TTACCAATTC CACCATATCA TACGGTAGGG TATTGGCAGC TTCCTTCAAG 2100  
30 GAATAGTTCT CTAAGTTATT TACATTTTGT CGTAATTTCT TGGCATACTT AGTTGTAATT 2160  
AATCGTTTTT CTTGCTATTC GAAAATCAAC TTGCGCTCCA GATAATAGCC TCTCAGCATT 2220  
TCATTGATAT TGTTGGGTTT GACACGATTG ATAACCCGTT CGACAAAGGC ACCACTGCTG 2280  
35 ATAATAGTTG TTTCTCGAAG ACGAGACTCC TGCATAAAAC TAATCAAAGA GCGTCTGTAG 2340  
ACTCCCTTCA GGTTTTCCAA ACTTTCAATA ATCATCTCCG TATTGGCAAG ATAGAGCTCT 2400  
40 GCAATTTGGT CATAATCAAG AGCACGGAGA CGGCTTTGCT CCTTGTCTCT CCAGCTACGG 2460  
AAGGTCTTTC CAAGAGTAAA AACTTCATGA AGGAGAAAAC GTAAAATCCT CAAGGAAACA 2520  
AGAAAATAAT AGGTCAGTCT TGAGGCAAGT TTACGATTGA TTCCTTGTTT TATATTTTTC 2580  
45 AGATAACGTT GGTAACTCG GTAAGCACGA TTGCTAATGT TCCCCTCTTC ATAGGCCTGT 2640  
TCCAAACCAT CACTTTCAAT ACTAAGAATC AAGAGTTTCA AAGCAGCCCA GTCTTCTTGA 2700  
50 TC 2702

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:  
55 (A) LENGTH: 6217 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- 60 (ii) MOLECULE TYPE: DNA (genomic)



-108-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

10	GAATTCCAAG AAGCTAGCCA AGAAAGTCGT GAACGTAGTG ATCCGCTAAA TAGTTATCTC	60
	CTTTTGTGAG GCTCCTTGAC GAAAGAAAAG CTTGCCGATA AATTAGGAGA TTTGGGTTAT	120
	AAAGCAAGTG CTGACCGAAA GATACCGCCC TATTTTCTTG CTTTTCGAAT ATTACTAAAT	180
15	CCCCTTATTT TAATTAGTTT AGCAATATTT GGCTTATCTT TCTTTGCTTT AGTGATTATC	240
	ACTCGGATTA AGGAAATGAG AGCAGCAGGT ATAAACTCT TTTCTGGTCA GACTCTCTTA	300
20	TCCATCATGG GGCATTCTTT ATCTACTGAT ATCAAATGGC TCCTTCTATC AGCCCTCCTT	360
	TCCTTCCTAG GTGGGGGTGT CGTTCTTTTT AGTCAAGGTT TGTTTTATCC TATCTTGTTA	420
	GCCACCTATG GTTTTGGGAT TAGTTTCTAT CTGTTGTTTT TATTGGCGAT TTCAATTTTA	480
25	CTAATGCTTC TTTATCTAAT GAGTTTGAAT ACAAAGCATT AGTCCCGTT ATTAGGGGGA	540
	GATCCCCCT GAACCCTGAT GAATAACCCA TTGTTTCAGT AGACCTGTTT TTTCAGTAGG	600
30	ATACGCTTTA AGACAGGTTG ACGTCTTACC AACGATTGAA AGAACTTGAA ATTTGAGACA	660
	AGATGGCAGG ATAGAGTAGA CTATTATCAC GATTTCCTTT GACTTAGGTT ATAGAGGTTG	720
	AGATTCAGAA AATCAGAGCA AGTGGTATGC CTTTACCAAG GGAGCAGCGA AGAAGAACAA	780
35	GCTCTTTATG TAAAGGATAA TCTGCTCCAT TTTGCCAATC CACAAGGAAA AAATGAACAG	840
	GGAGAGACAC TGGATACCTA TAGTCCAGAT GCTAATACGC TCTATGTTAG TCCCAGTTAT	900
40	TTGGACAAGG AAAAGGTCGT GGTAGATGCT GAGACCAAAC AGAAGTTAGC CCATCTCCAA	960
	AAAGGTGAGT TTATCCTCTT GCTCCCAGAA CATTGCGCT CTCGAGAAGC AGAACTTAAG	1020
	AAAGTTTTTG AAGAAAGATT GAGTTATTAT GGAAATCTG GTGAGGAGGC AAGTGCTCCT	1080
45	TTGGATTATG AGATGAAAGC GCACGTTAGT TATCTTTCAA TGGGAGAAAA GCGGTTTGTT	1140
	TATAATAACG GTGAGAATCC CGTATCTACT CAGTATTTGA CTGATCCGAT TTTAGTTGTA	1200
50	TTACGCCCGA CTTCTACAGG TGATAGTTTT ATATCCTTAT CTAGTTGGTC TATCAATGCT	1260
	GGAAAACAAC TCTTTATCAA AGGATATGAG AGTGGGCTAG AACTCTTGAA GAAAGCTGGA	1320
	ATTTATGAGC AAGTATCCTA TCTTAAAGAA GGAAGAAGTG TTTATCTAAC TCGTTATAAT	1380
55	GAAGTTCAAA CTGAAACAGC AACTTTAATC TTAGGAGCTA TTGTGGGGAT AGCTAGTTCC	1440
	TTGTTACTCT TTTATTCTGT CAATCTTCTA TATTTGAGC AATTCCGCCG AGATATCTTG	1500
60	ATTAAACGAA TTTCAGGTTT ACGATTTTTT GAAACACATG CTCAGTATAT GGTTAGTCAA	1560

-109-

	TTTGCCAGTT TTGTATTTGG TGCTAGTCTC TTTATTTTAA GCAGTCGAGA CTTGGTGATT	1620
	GGCTTGCTCA CTTTATTAGT CTTTCTAGCT AGTGCAGTTT TGACGCTTTA CCGTCAAGCG	1680
5	CAGAAAGAAT CTCGTGTTTC TATGACAATT ATGAAAGGAA AATAGGATGA TTGAACTAAA	1740
	GAATATATCT AAAAAATTGG AAGCCGTCAG CTATTTTCAG ATACGAATCT TCATTTTGAA	1800
	GGTGGGAAAA TTTATGCCTT AATCGGTACA AGTGGCTGTG GTAAGACAAC ACTCTCGAAT	1860
10	ATGATTGGAC GATTGGCGCC ATATGACAAA GGGCAAATCA TCTATGATGG CACTTCTCTT	1920
	AAGGACATCA AGCCTTCTGT TTTCTTTAGA GATTACTTAG GATACTTATT TCAAGATTTT	1980
15	GGCTTAATTG AAAGCCAAAC CGTCAAAGAG AATCTCAATC TGGGTTTAGT TGGTAAAAAG	2040
	TTGAAGGAAA AAGAGAAAAT CTCTTTGATG AAACAAGCTC TAAACCGTGT AAACCTCTCT	2100
	TATTTGGATT TGAAGCAACC TATATTTGAG TTATCAGGAG GAGAAGCACA ACGTGTTGCA	2160
20	CTAGCGAAGA TAATTTTAAA GGATCCGCCT TTGATTCTTG CAGATGAACC AACCGCTTCC	2220
	TTAGACCCCA AAAATTCTGA GGAATTACTT TCCATCCTAG AATCTTTAAA AAATCCGAAT	2280
25	CGGACCATTA TTATTGCGAC CCACAATCCT CTGATTTGGG AGCAAGTGGA TCAGGTCATT	2340
	CGAGTTACCG ATTTATCACA TAGATGATAT GGTAAGATTC AGTTAGAAGA AAGAGTCACA	2400
	AACACACTTT GTGGCTTTTT TATTTCCATA AAAATGGTAA AATAGTAGGA GTAGAAATGA	2460
30	GTTTCGAGACA TGAAAGTAAT AGATCAATTT AAAAATAAGA AAGTTCCTGT TTTAGGTTTG	2520
	GCCAAGTCTG GTGAATCTGC AGCTCGTTTG TTGGACAAGC TAGGTGCCAT TGTGACAGTA	2580
35	AATGATGGGA AACCTTTCGA GGACAATCCA GCTGCCCAA GTTTGCTGGA AGAAGGGATC	2640
	AAGGTCATTA CAGGTGGCCA TCCTTTGGAA CTCTTGATG AAGAGTTTGC CCTTATGGTG	2700
	AAAAATCCAG GTATCCCCTA CAACAATCCC ATGATTGAAA AGGCTTTGGC CAAGAGAATT	2760
40	CCAGTCTTGA CTGAGGTGGA ATTGGCTTAT TTGATTTTCA AAGCACCGAT TATTGGTATC	2820
	ACAGGATCGA ACGGTAAGAC AACCACAACG ACTATGATTG GGAAGTTTT GACTGCTGCT	2880
45	GGGCAACATG GTCTTTTATC AGGGAATATC GGCTATCCTG CCAGTCAGGT TGCTCAAATA	2940
	GCATCAGATA AGGACACGCT TGTATGGAA CTTTCTTCTT TCCAACATCAT GGGTGTTCAA	3000
	GAATTCATC CAGAGATTGC GGTATTACC AACCTCATGC CAACTCATAT CGACTACCAT	3060
50	GGGTCATTTT CTGAATATGT AGCAGCCAAG TGGAATATCC AGAACAAGAT GACAGCAGCT	3120
	GATTCCTTG TCTTGAACCT TAATCAAGAC TTGGCAAAAG ACTTGACTTC CAAGACAGAA	3180
55	GCCACTGTTG TACCATTTTC AACACTTGAA AAGGTTGATG GAGCTTATCT GGAAGATGGT	3240
	CAACTCTACT TCCGTGGTGA AGTAGTCATG GCAGCGAATG AAATCGGTGT TCCAGGTAGC	3300
	CACAATGTGG AAAATGCCCT TGCAGCTATT GCTGTAGCCA AGCTTCGTGA TGTGGACAAT	3360
60	CAAACCATCA AGGAACTCT TTCAGCCTTC GGTGGTGTCA AACACCGTCT CCAGTTTGTG	3420

-110-

	GATGACATCA AGGGTGTAA ATTCTATAAC GACAGTAAAT CAACTAATAT CTTGGCTACT	3480
5	CAAAAAGCCT TATCAGGATT TGACAACAGC AAGGTCGTCT TGATTGCAGG TGGTTTGGAC	3540
	CGTGGCAATG AGTTTGACGA ATTGGTGCCA GACATTACTG GACTCAAGAA GATGGTCATC	3600
	CTGGGTCAAT CTGCAGAACG TGTCAAACGG GCAGCAGACA AGGCTGGTGT CGCTTATGTG	3660
10	GAGGCGACAG ATATTGCAGA TGCGACCCGC AAGGCCTATG AGCTTGCGAC TCAAGGAGAT	3720
	GTGGTTCTTC TTAGTCCTGC CAATGCCAGC TGGGATATGT ATGCTAACTT TGAAGTACGT	3780
15	GGCGACCTCT TTATCGACAC AGTAGCGGAG TTAAAAGAAT AAAATATGAA AAAAATTGTC	3840
	TTTACAGGTG GGGGGACGGT TGGACACGTT ACCCTCAATC TTTTGTTAAT GCCCAAGTTC	3900
	ATCGAAGATG GTTGGGAAGT CCACTATATC GGGGACAAGC GTGGTATCGA ACACCAAGAA	3960
20	ATCCTTAAGT CAGGTTTGGA TGTCACTTTC CACTCCATTG CGACTGGGAA ATTGCGTCGC	4020
	TATTTCTCTT GGCAAAATAT GCTGGACGTC TTCAAAGTTG GCTGGGGAAT CGTCCAATCG	4080
25	CTCTTTATCA TGTTGCGACT TCGTCCACAG ACCCTTTTTT CAAAGGGGGG CTTTGTCTCA	4140
	GTACCGCCTG TTATCGCAGC GCGTGTGTCA GGAGTGCCTG TCTTTATTCA CGAATCTGAC	4200
	CTGTCTATGG GCTTGCCAA TAAAATCGCC TATAAATTG CGACTAAGAT GTATTCAACC	4260
30	TTTGAGCAAG CTTTCAGTTT GTCTAAGGTT GAGCATGTGG GAGCAGTGAC CAAGGTTTCA	4320
	GATCAAAAAA ATCCAGAACC AGATGAATTG GTGGATATTC AAACCCACTT TAATCATAAA	4380
35	TTGCCGACTG TATTGTTTGT TGGCGGTTCT GCAGGTGCTC GTGTCTTTAA CCAATTGGTG	4440
	ACAGACCATA AGAAAGAACT AACAGAGCGC TACAATATTA TCAATCTAAC TGGAGATTCT	4500
	AGTCTGAACG AGTTGAGCCA AAATCTTTTT CGTGTGACT ATGTGACCGA TCTCTATCAA	4560
40	CCCTTGATGG AATTGGCTGA TATTGTTGTG ACACGAGGTG GTGCCAATAC GATTTTGTAG	4620
	CTCTTGCGA TAGCAAAATT GCATGTCATT GTGCCGCTTG GTCGTGAAGC TAGTCGTGGT	4680
45	GACCAGATTG AAAATGCAGC TTAATTTGTA AAAAAAGGCT ATGCAGAAGA CCTTCAAGAA	4740
	AGCGATTGA CCTTGGATAG TTTGGAAGAG AAGCTTCTC ACTTACTAAG TCACAAGGAA	4800
	GATTACCAAG CTAAGATGAA AGCTTCTAAG GAATTGAAAT CTCTAGCAGA TTTTATCAA	4860
50	TTGTTGAAA AAGATTTATC ATAAGGAAAG TAAATGTCAA AAGATAAGAA AAATGAGGAC	4920
	AAAGAAACCC TCGAAGAATT GAAAGAGTTA TCAGAATGGC AGAAACGAAA CCAAGAATAT	4980
55	CTAAAAAGA AGGCTGAAGA AGAGGTGGCT CTAGCTGAGG AGAAGGAAA GGAAAGACAA	5040
	GCTCGAATGG GAGAAGAATC TGAGAAGTCA GAGGACAAAC AGGACCAGGA GAGTGAAACA	5100
	GACCAGGAAG ATTCAGAATC AGCTAAGGAA GAGTCTGAAG AAAAAGTAGC ATCCTCAGAG	5160
60	GCTGACAAAG AGAAAGAAGA ACCAGAGTCT AAAGAGAAGG AGGAACAGGA TAAAAAGCTT	5220

-111-

GCTAAAAAGG CTACAAAGGA AAAACCAGCC AAAGCAAAGA TTCCTGGTAT CCATATCTTG 5280  
CGAGCCTTCA CGATTTTATT TCCAAGTCTG CTTTATTGA TTGTCTCTGC CTACTTGCTC 5340  
5 AGTCCTTATG CGACCATGAA AGATATTCGT GTTGAGGGAA CGGTGCAAAC TACAGCTGAT 5400  
GATATTCGAC AGGCTTCAGG CATTCAAGAT TCGGATTATA CGATTAACCT TCTGCTAGAC 5460  
AAGGCAAAAT ATGAAAAGCA GATTAAGTCT AACTATTGGG TTGAATCAGC TCAACTTGTC 5520  
10 TATCAATTTT CAACTAAGTT CACTATTAAG GTCAAGGAAT ATGATATTGT GGCCTACTAT 5580  
ATTTCTGGTG AAAATCATTA TCCTATTCTT TCCAGTGGTC AGCTTGAGAC TAGTTCTGTG 5640  
15 AGTCTGAACA GTTTACCAGA AACTATTTTA TCAGTTCTCT TTAATGATAG TGAACAAATC 5700  
AAGGTTTTTG TCTCAGAACT TGCTCAAATT AGCCCAGAAC TCAAGGCGGC TATCCAAAAG 5760  
GTGGAATTAG CCCCAGCAA GGTGACATCC GATTTAATTC GATTGACCAT GAATGATTCTG 5820  
20 GACGAACTCT TGGTTCCTCT ATCTGAAATG AGTAAGAAAC TGCCATATTA CAGTAAGATT 5880  
AAGCCACAAT TGTCAGAACC GAGTGTGGTC GACATGGAAG CTGGAATTTA CAGTTACACT 5940  
25 GTGGCGGATA AATTAATTAT GGAGGCTGAG GAAAAAGCCA AACAAGAGGC CAAGGAAGCT 6000  
GAGAAAAAAC AAGAAGAAGA ACAGAAAAAA CAAGAGGAAG AGAGCAATCG AAATCAAACA 6060  
AATCAGCGTT CATCGCGTCG CTAGGTTTAC CTTTCTCTT ATAGTTCTTT AGTGACCATG 6120  
30 TTTTACGTT TAATATTTGA CATTGTTTT TCTTATGTT ACATCTGCAA TGTAATCGAT 6180  
TACAAAATAA TTTTGATGA AGAAGGTAAC ACATATG 6217

35 (2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1491 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

45 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CTTGACACTT GATTGCGACT GTTGAATCTT ATCTCTCCAA GAAAAACACG TGAAGATGTT 60  
55 GAGTCTGCTG TCAGCAAGCT TGAAAGTAGC ACATCTGAGA AACATTGGAT CCATCTGCAG 120  
TTTCTCGTGG GTCTAGCTTG GATCGTGATG ACAATGGTCT TTTGACTCTT GCTGGCGGTA 180  
60 AAATCACAGA CTACCGTAAG ATGGGTGACG AGGCGCTATG GAGCGCGTGG TTGACATCCT 240

-112-

CAAAGCAGAA TTTGACCGTA GCTTTAAATT GATCAATTCT AAAACTTACC CTGTTTCAGG 300  
TGGAGAATTG AACCCAGCAA ATGTGGATTC AGAAATCGAA GCCTTTGCGC AACTTGGAGT 360  
5 TTCACGTGGT TTGGATAGCA AGGAAGCTCA TTACCTAGCA AATCTTTACG GTTCAAATGC 420  
ACCGAAGGTC TTTGCACTTG CTCACAGCTT GGAACAAGCG CCAGGACTCA GCTTGGCAGA 480  
TACTTTGTCC CTTCACTATG CAATGCGCAA CGAGTTGGCT CTTAGCCCAG TTGACTTCCT 540  
10 TCTTCGTCGT ACCAACCATA TGCTCTTTAT GCGTGATAGC TTGGATAGCA TCGTTGAGCC 600  
AGTTTTGGAT GAAATGGGAC GATTCTATGA CTGGACAGAA GAAGAAAAAG CAACTTACCG 660  
15 TGCTGATGTC GAAGCAGCTC TCGCTAACAA CGATTTAGCA GAATTAAAAA ATTAAGAAAA 720  
AATAAAGAG GTGGAGGGCA GCATTCTTGG TCGCCCGTCC CTTCTTTTGA ATGGAGACAG 780  
AAAGATGATG AATGAATTAT TTGGAGAATT TCTAGGGACT TTAATCCTGA TTCTTCTAGG 840  
20 AAATGGTGTT GTTGCAAGTG TGTTCTTCC TAAACCAAG AGCAATAGCT CAGGTTGGAT 900  
TGTGATTACT ATGGGTTGGG GGATTGCAGT TGCGGTTGCA GTCTTTGTAT CTGGCAAGCT 960  
25 CAGTCCAGCT CATTTAAACC CAGCTGTGAC CATCGGTGTG GCCTTAAAAG GTGGTTTGCC 1020  
TTGGGCTTCC GTTTTGCCTT ATATCTTAGC CCAGTTCGCA GGGGCCATGC TGGGTCAGAT 1080  
TTTGGTTTGG TTGCAATTCA AACCTCACTA TGAGGCAGAA GAAAATGCAG GCAATATCCT 1140  
30 GGCAACCTTC AGTACTGGAC CAGCCATCAA GGATACTGTA TCAAACCTGA TTAGCGAAAT 1200  
CCTTGGAACC TTTGTTTTGG TGTGACAAT CTTTGCTTTG GGTCTTTACG ATTTTCAGGC 1260  
35 AGGTATCGGA ACCTTTGCAG TGGGAACTTT GATTGTGCGT ATCGGTCTAT CACTAGGTGG 1320  
GACAACAGGT TATGCCTTGA ACCCAGCTCG TGACCTTGGA CCTCGTATCA TGCACAGCAT 1380  
CTTGCCAATT CCAAACAAGG GAGACGGAGA CTGGTCTTAC GCTTGGATTC CTGTTGTAGG 1440  
40 CCCTGTTATC GGAGCAGCCT TGCCCGTGCT TGTATTGTCA CTTTCTAAT C 1491

(2) INFORMATION FOR SEQ ID NO:54:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1229 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
50 (ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
55 (iv) ANTI-SENSE: NO  
60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

-113-

ACAACGGATA ATGTCATCGA TCTCTTTGAA CACATCTTTA AGGAATGTTT AACGAAAACA 60  
TTGTGATGGC GGGCAAGGTC AATCTCTTGA ATTTTGCCAA TCTAGCAGCC TATCAGTTCT 120  
5 TTGACCAACC GCAAAGGTG GCCTTGGAGA TTCGTGAGGG GTTGCCTGAG GATCAGATGC 180  
AAAATGTTCTG TGTTCGAGAC GGTCAAGAGT CCTGTTTAGC TGACCTAGCG GTGATTAGTA 240  
10 GTAAGTTCCT CATTCCTTAT CGGGGAGTTG GAATTCTAGC CATTATCGGT CCAGTTAATC 300  
TGGATTACCA ACAGCTAATC AATCAAATCA ATGTGGTCAA CCGTGTTTTG ACCATGAAGT 360  
TGACAGATTT TTACCGCTAC CTCAGCAGTA ATCATTACGA AGTACATTAA GATTGAAATC 420  
15 ATTAAAGGAG GCGAACATGG CCCAAGATAT AAAAAATGAA GAAGTAGAAG AAGTTCAAGA 480  
AGAGGAAGTT GTGGAACAG CTGAAGAAAC AACTCCTGAA AAGTCTGAGT TGGACTTGGC 540  
AAATGAACGT GCAGATGAGT TCGAAAACAA ATATCTTCGC GCTCATGCAG AAATGCAAAA 600  
20 TATCCAACGC CGTGCCAATG AAGAACGTCA AAAGTTGCAA CGTTATCGTA GCCAGGACTT 660  
GGCAAAAGCA ATCTTACCAT CTCTTGACAA CCTTGAGCGT GCACTTGCAG TTGAAGGTTT 720  
25 GACAGATGAT GTGAAGAAGG GCTTGGCGAT GGTGCAAGAA AGCTTGATTC ACGCTTTGAA 780  
AGAAGAAGGA ATTGAAGAAA TCGCAGCAGA TGGCGAATTT GACCATAACT ACCATATGGC 840  
CATCCAACT CTCCCAGGAG ACGATGAACA CCCAGTAGAT ACCATCGCCC AAGTCTTTCA 900  
30 AAAAGGCTAC AAAGTCCATG ACCGCATCCT ACGCCAGCA ATGGTAGTGG TGTATAACTA 960  
AGATACAAAC GCTCGTAAAA AGCTCGCAGT AAAAATAGGA GATTGACGAG TGTTCGATGA 1020  
35 ACACAAGAAA ATCTATCTTT TTTACTCAGA GCTTAGGGCG TGTTCGATTC GGCAATTCTG 1080  
ACGGTAGCTA AAGCAACTCG TCAGAAAACG GCAATCGCTA TGACGTTTGC CTAGCTTCCT 1140  
TACTAACTCG TCGTCGAAAT AAAATCGATT TCGACTCCTC GTGTCGCAAT TTACATAATA 1200  
40 GAAAACCTGT CCGAACGACA TAACTATG 1229

## (2) INFORMATION FOR SEQ ID NO:55:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5816 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
50  
(ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
55 (iv) ANTI-SENSE: NO  
60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

-114-

	AAGAAGAAGA CTGTATGGAT AATCGACCAA TTGGTTTTTT GGATTCGGGT GTCGGGGGCT	60
	TGACCGTTGT GCGCGAGCTC ATGCGCCAGC TTCCCCATGA AGAAATCGTC TATATTGGAG	120
5	ATTGCGCGCG GCGGCCCTAT GGGCCCCGTC CTGCTGAGCA AATTCGTGAA TATACTTGGC	180
	AGCTGGTCAA CTTTCTCTTG ACCAAGGATG TCAAAATGAT TGTCAATGCT TGTAACACTG	240
10	CGACTGCGGT TGTCTGGGAA GAAATCAAGG CTCAACTAGA TATTCCTGTC TTGGGTGTAA	300
	TTTTGCCAGG AGCTTCGGCA GCCATCAAGT CCAGTCAAGG TGGGAAAATC GGAGTGATTG	360
	GAACGCCCAT GACGGTACAA TCAGACATAT ACCGTCAGAA AATCCATGAT CTGGATCCCG	420
15	ACTTACAGGT GGAGAGCTTG GCCTGTCCCA AGTTTGCTCC CTTGGTTGAG TCAGGTGCCC	480
	TGTCAACCAG TGTACCAAG AAGGTGGTCT ATGAAACCCT GCGTCCCTTG GTTGGAAGG	540
20	TGGATAGCCT GATTTTGGGC TGTACTCATT ATCCACTCCT TCGCCCTATT ATCCAAAATG	600
	TGATGGGGCC AAAGGTTGAG CTCATCGATA GTGGGGCAGA GTGCGTACGG GATATTTGAG	660
	TCTTACTCAA TTATTTTGAA ATCAATCGTG GTCGCGATGC TGGACCACTC CATCACCGTT	720
25	TTTACACAAC AGCCAGTAGC CAAAGTTTTG CACAAATTGG TGAAGAATGG CTGGAAAAAG	780
	AGATTCAATG GGAGCATGTA GAATTATGAC AAATAAAATT TATGAATATA AGGATGACCA	840
30	GAAGTGGTAT GTTGGGTCTT ATAGTATTTT TGGTGGCGTT AACAGTTTGA GCGACTATAA	900
	GGCAGATTTT CCTCTGTTTG AATTCTCCAA AATATTTGGA GATGAAGAGT ATGGTTTCCC	960
	GCTTTCAGTT ACTGTTTAC GCTATGGTTC TACCTACCGT TTGTTCTCCT TTGTGGTAGA	1020
35	CATGCTTAAT CAAGAAATGG GACGAACTT GGAAGTTATT CAACGTCATG GGGCCCTGCT	1080
	CTTGGTTGAA AATGGGCAAC TCTGTATGT AGAATTGCCT AAAGAAGGGG TCAATGTTCA	1140
40	TGATTTCTTT GAGACAAGCA AGGTCAGAGA AACCTTGTTG ATTGCGACTC GTAACGAAGG	1200
	TAAAACCAAG GAATCCGAG CTATCTTTGA TAAGTTAGGC TACGATGTGG AAAATCTTAA	1260
	TGACTACCTT GACCTGCCTG AAGTAGCAGA AACAGGTATG ACCTTTGAAG AAAATGCCCG	1320
45	CCTTAAGGCA GAAACCATT CTCAATTAAC GGGCAAGATG GTTTTGGCAG ATGATTCTGG	1380
	TCTCAAAGTC GATGTCCTTG GTGGCTTACC AGGCGTCTGG TCAGCTCGTT TCGCAGGTGT	1440
50	GGGAGCAACT GACCGTGAAA ATAATGCCAA ACTCTTGAC GAATTGGCCA TGGTCTTTGA	1500
	ACTCAAGGAC CGCTCGGCTC AGTTCCACAC AACCCTAGTC GTAGCCAGCC CAAATAAGGA	1560
	AAGTTTAGTT GTTGAAGCAG ACTGGTCAGG TTATATTAAC TTTGAACCTA AGGGTGAAAA	1620
55	TGGCTTTGGC TATGATCCCC TCTTCCTTGT AGGAGAGACA GGTGAGTCAT CAGCTGAATT	1680
	AACCCTGGAA GAAAAAATA GTCAATCTCA CCGTGCCTTA GCCGTTAAGA AACTTTTGGG	1740
60	GGTATTTCCA TCATGGCAAA GCAAACCATC ATTGTAATGA GCGATTCCCA TGGCGATAGC	1800
	TTGATTGTGG AAGAAGTCCG TGATCGCTAT GTGGGCAAAG TCGATGCCGT TTTTCATAAC	1860

-115-

	GGCGATTCTG AACTACGTCC GGATTCTCCA CTTTGGGAGG GCATCCGCGT TGTAAAGGG	1920
	AACATGGACT TCTACGCCGG CTACCCAGAA CGTCTGGTGA CTGAGCTTGG TTCGACCAAG	1980
5	ATTATCCAAA CTCATGGTCA CTTGTTTGAC ATCAATTTCA ACTTTCAAAA GTTGGACTAC	2040
	TGGGCTCAGG AGGAAGAGGC CGCTATCTGC CTCTATGGTC ACTTGCAATG GCCAAGTGCT	2100
10	TGGATGGAAG GCAAGATCCT CTTTCTAAAT CCAGGTTCTA TCAGTCAACC ACGAGGTACC	2160
	ATCAGAGAAT GTCTCTATGC TCGTGTGGAG ATTGATGATA GTTACTTCAA AGTGGACTTT	2220
	TTGACACGAG ATCACGAGGT GTATCCAGGT TTGTCCAAGG AGTTTAGCCG ATGATTGCCA	2280
15	AGGAGTTTGA GACTTTCTTG TTGGGGCAGG AGGAAACTTT TTTGACCCCT GCTAAAAATC	2340
	TAGCTGTGTT GATTGATACC CACAATGCGG ATCATGCGAC CCTCTTGCTC AGTCAGATGA	2400
20	CCTATACCCG TGTTCCTGTT GTGACAGATG AAAAACAGTT TGTGGGACG ATTGGACTCA	2460
	GAGATATTAT GGCTTATCAG ATGGAGCATG ACTTGAGCCA AGAAATCATG GCGGATACGG	2520
	ATATCGTTCA TATGACAAAA ACGGACGTAG CGGTTGTTTC GCCTGATTTC ACCATTACGG	2580
25	AGGTCTTGCA CAAGCTAGTA GATGAGTCTT TCTTACCGGT CGTGGATGCA GAGGGTATTT	2640
	TCCAAGGGAT TATTACGCGC AAGTCCATCC TCAAGGCCGT TAATGCCCTC TTGCATGACT	2700
30	TTAGTAAGGA ATATGAGATT CGATGCCAAT GAGAGACAGG ATTTACAGCT TTTTAGAGGA	2760
	AAAGCAGGGC TTGTCTGTCA ATTCCAAGCA GTCCTATAAG TATGATTTGG AGCAATTTTT	2820
	AGACATGGTA GGTGAGCGGA TTTCTGAGAC CAGTCTCAAG ATTTACCAAG CCCAGCTAGC	2880
35	CAATCTAAAA ATCAGCGCCC AGAAGCGAAA GATTTGCGCC TGTAACCAAT TTCTATACTT	2940
	TCTCTATCAA AAAGGAGAGG TGGACAGCTT TTACCGCTTG GAATTAGCTA AACAAGCTGA	3000
40	AAAGAAGACG GAAAAGCCAG AGATTCTATA CCTAGACTCT TTTTGGCAGG AAAGCGACCA	3060
	TCCAGAGGGC CGCTTGCTAG CGCTCTTAAT CCTAGAAATG GGGCTCTTGC CCAGTGAGAT	3120
	TTTAGCCATC AAGGTTGCGG ACATCAATCT GGATTTTCAG GTGCTGCGAA TCAGCAAGGC	3180
45	TTCCCAACAG AGGATTGTCA CCATTCCAC GGCCTTGCTT TCAGAATTGG AACCCTTGAT	3240
	GGGGCAGACC TATCTTTTTG AAAGAGGAGA GAAACCCTAT TCTCGTCAGT GGGCCTTTTCG	3300
50	TCAGTTAGAA TCTTTTGTCG GGAGAAGGTT TCCATCCTTA TCAGCTCAAG TCTTACGTGA	3360
	CAGTTTATTC TAAGCAAATA GAAACAGGTC GATTTGTACG AATTGCAAAA AATTAGGATT	3420
	AAAAACAGTC CTGACCTTAG AAAATATAGA TAATGGATAT TAAATTAAAA AGATTTTTGA	3480
55	AGGACCCTGG ACTTGCTCTT TGCATCTGGT TTCTAAGTAC CAAGATGGAT ATCTACGATG	3540
	TGCCCATTAC GGAAGTCATC GAACAGTATC TAGCCTATGT TTCAACCCTG CAGGCCATGC	3600
60	GTCTGGAAGT GACGGGTGAG TACATGGTCA TGGCTAGTCA GCTCATGCTG ATTAAGAGTC	3660



-116-

	GTAAACTCCT	TCCGAAGGTA	GCAGAAGTGA	CAGACTTGGG	GGATGACCTG	GAGCAGGACC	3720
	TCCTCTCTCA	AATCGAAGAA	TATCGCAAGT	TCAAGCTCTT	GGGTGAGCAC	TTGGAAGCCA	3780
5	AGCACCAAGA	ACGGGCCAG	TATTATTCCA	AAGCGCCGAC	AGAGTTGATT	TACGAAGATG	3840
	CGGAGCTTGT	GCATGACAAG	ACGACCATTG	ACCTCTTTTT	GGCTTTTTCA	AATATCCTAG	3900
10	CCAAGAAAA	AGAGGAGTTT	GCACAAAATC	ACACGACGAT	CTTGCGGGAT	GAGTATAAGA	3960
	TTGAGGACAT	GATGATTATT	GTGAAAGAAT	CCTTGATTGG	ACGAGATCAA	TTGCGCTTGC	4020
	AGGATTGTGT	CAAGGAAGCC	CAGAATGTCC	AAGAGGTCAT	CACCCTCTTT	TTGGCAACCC	4080
15	TAGAGTTAAT	CAAAACCCAG	GAGCTGATCC	TCGTGCAAGA	GGAGAGTTTC	GGAGATATCT	4140
	ATCTCATGGA	AAAGAAGGAA	GAAAGTCAAG	TGCCTCAAAG	CTAGACTTGA	TAGAGAGGAA	4200
	AGATGAGTAC	TTTAGCAAAA	ATAGAAGCGC	TCTTGTTTGT	AGCGGGTGAA	GATGGGATTC	4260
20	GGGTCGCCCA	GTTAGCTGAA	CTCCTCTCTC	TGCCACCGAC	AGGCATCCAG	CAGAGTTTAG	4320
	GAAAATTAGC	CCAGAAGTAT	GAAAAGGACC	CAGATTCCAG	TTTGGCTTTG	ATTGAGACAA	4380
25	GTGGTGCTTA	TAGATTGGTG	ACCAAGCCTC	AATTTGCAGA	GATTTTGAAG	GAATACTCTA	4440
	AGGCGCCTAT	CAACCAGAGC	TTGTCTCGGG	CTGCCCTTGA	GACCTTGTCC	ATTATTGCCT	4500
	ACAAACAGCC	GATTACGCGG	ATAGAAATTG	ATGCCATCCG	TGGGGTTAAC	TCGAGTGGAG	4560
30	CCTTGGCAAA	GTTGCAAGCT	TTTGACCTGA	TAAAGGAAGA	CGGGAAAAAG	GAAGTATTGG	4620
	GGCGCCCCAA	CCTCTATGTG	ACTACGGATT	ATTTCCCTAGA	TTACATGGGG	ATAAACCATT	4680
35	TAGAAGAATT	ACCAGTGATT	GATGAGCTTG	AGATTCAAGC	CCAAGAAAGC	CAATTATTTG	4740
	GTGAAAGGAT	AGAAGAAGAT	GAGAATCAAT	AAGTATATTG	CCCACGCAGG	TGTGGCCAGT	4800
	AGGAGAAAAG	CAGAAGAGCT	GATCAAGCAA	GGTTTGGTGA	CAGTTAACGG	ACAAGTGGTG	4860
40	CGTGAAGTAG	CAACCACTAT	CAAGTCAGGC	GACAAGGTCG	AAGTTGAAGG	TCAACCTATC	4920
	TACAACGAAG	AAAAGGTCTA	TTATCTGCTT	AACAAACCAC	GCGGTGTCAT	TTCCAGTGTA	4980
45	ACAGATGACA	AGGGTCGCAA	GACGGTTGTC	GACCTCTTGC	CCAATGTCAA	AGAGCGCATT	5040
	TACCCTGTGG	GTCGTTTGGA	CTGGGATACA	TCAGGAGTCT	TGATTTTGAC	CAATGATGGG	5100
	GACTTTACAG	ACGAGATGAT	TCACCCTCGT	AATGAGATTG	ACAAGGTTTA	TGTCGCGCGT	5160
50	GTTAAAGGTG	TGGCCAATAA	GGACAATCTT	CGCCCCTTGA	CCCGTGGTCT	TGAGATTGAT	5220
	GGTAAGAAAA	CCAAGCCAGC	TGTTTATGAA	ATTCTCAAAG	TGGACCCAGT	CAAAAATCGC	5280
55	TCTGTGGTGC	AGTTGACCAT	CCATGAAGGG	CGTAACCATC	AGGTTAAAAA	GATGTTTGAA	5340
	GCTGTTGGTC	TCCAAGTAGA	TAAGTTGTCT	CGGACTCGTT	TCGGACACCT	AGACTTGACA	5400
	GACTCCGTCC	AGGAGAATCC	CGTCGTCTTA	ATAAAAAAGA	AATCAGCCAA	CTACACACCA	5460
60	TGGCTGTAAC	TAAGAAATAA	TGAAACGAAT	TTTAATAGCG	CTTGTGCGCT	TTTACCAACG	5520

-117-

TTTTATCTCA CCAGTCTTTC CACCCTCTTG TCGCTTTGAG CTGACTTGTT CCAACTACAT 5580  
GATTCAGGCT ATTGAAAAAC ATGGTTTTAA GGGGGTATTG ATGGGCTTGG CTCGGATTTT 5640  
5 ACGTTGTCAT CCCTGGTCGA AAACAGGTAA GGACCCCGTT CCAGACCACT TTTCCCTTAA 5700  
ACGAAATCAA GAAGGGGAAT GAGGTGGGGT AAATAGATTT CAAAATGATA AAAACGCATC 5760  
10 CTATCAGGTT TGAGTGAAC TATAGGATG CGTTTTAGAA TGTCAAAATT TTATAC 5816

## (2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:  
15 (A) LENGTH: 725 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

30 TTGAAAATAA TTATGAACCG CAATATATTA ATATCCGAGG AAAAGGCCCT CTTATCAATG 60  
ACTTGAAAAA AGAAGCTAAA AAAGCTAATA AAGTTTTTCT CGCGAGTGAC CCGGACCGTG 120  
35 AAGGAGAAGC GATTTCTTGG CATTGGGCCC ATATTCTCAA CTTGGATGAA AATGATGCCA 180  
ACCGTGTGGT CTTCAATGAA ATCACCAAGG ATGCAGTCAA AAATGCTTTT AAAGAACCTC 240  
GTAAGATCGA TATGGACTTG GTCGATGCCC AACAAGCTCG TCGGATCTTG GATCGCTTGG 300  
40 TAGGGTATTC GATTTGCGCT ATTTTGTGGA AGAAGGTCAA GAAGGGCTTG TCAGCAGGTC 360  
GCGTTCAGTC CATTGCCCTT AAATCATCA TTGACCGTGA AAATGAAATC AATGCCTTCC 420  
45 AGCCAGAAGA ATACTGGACA GTTGATGCTG TCTTTAAAAA GGGAACCAA CAATTTTCATG 480  
CTTCCTTCTA TGGAGTAGAT GGTA AAAAGA TGAAGTAC CAGCAATAAC GAAGTCAAGG 540  
AAGTCTTCTC TCGTCTGACG AGTAAAGACT TTTCAGTAGA TCAGGTGGAT AAGAAAGAGC 600  
50 GTAAGGCAAA TGCTCCTTTA CCCTATACCA CTTTCATCTAT GCAGATGGGA TGCTGCCAAT 660  
AAAATCAATT TCCGTACTCG AAAAACCATG ATGGTTGCCC AACAAGCTCT ATGAAGGAAT 720  
55 TATAT 725

## (2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:  
60 (A) LENGTH: 1935 base pairs  
(B) TYPE: nucleic acid

-118-

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

15	AACCCATCTG AAGAACTTTT CCGTGCTGCT AGCTCAGCTA TCGATAAAGC AGAAACTAAA	60
	GGTTTGATTG ACAAAAACAA AGCAAGCCGC GATAAAGCTC GTCTTTCAGC TAAACTTGCT	120
	AAATAAGAAA CAGTCCATAG AGGCTGTTTT TTTGTCTCCA AATAGGAAAA GGTAGAAAAT	180
20	GAAAATCACA ATTATCGGAT ATTCTGGTTC TGGTAAGTCA ACTCTAGCAG AAAAGTTATC	240
	TAATACTAC TCCATTCCAA AACTGCACAT GGACACACTC CAATTTCAAC CTGGTTGGCA	300
25	AGACAGTGAC TGCGAATGGA TGTTAACCGA GATAAAAAAC TTTCTCACCA AGCATAAAGC	360
	TTGGGTCATC GATGGTAATT ATTCTTGGTG CTAATACCAA GAACGAATGC AAGAAGCTGA	420
	CCAAATCATC TTTCTCAATT TTTTGCCATT GACCTGTCTC TTTAGAGCCT TTAAGCGTTA	480
30	TCTTAAATAC CGTGGAAG TCAGAGAAAG TATGGCGGCA GATTGCCCTG AACGCTTTGA	540
	GTGGGAGTTT ATCAGATGGA TTCTTTGGGA TGGGCGTAGC AAAACTCAAA AAGAAAATTA	600
35	CCAAAACTT TGCCAAGAAT ATTACATAA AGTCACTATC CTTGAAATC AGAGAGAGCT	660
	AGATCAATTT CTGGATAAGA AAAGGAAGTC CTACAATTCA TAAAGGGCTT CCTTTTTGGC	720
	TATAATTATT CTGCAATCAA GGTTCCTCAA CCAACCTTCA TCATATCAGT GAAGGTATTT	780
40	TGACGTTCTT CTGCAGTTGT GTCTTCGTCT GGATTGACCA AGCTATCAGA GATGGTCATG	840
	ATAGCTAGCG CATCAACATG GTATTGGGCA GCAAGATAGT AAAGAGCTGC TGCTTCCATT	900
45	TCCACAGCCT TGAATCCCA TTTACCAAGC TCGATATTCT TTTCAAAGTA ATTTGAGTAA	960
	AAGACATCAG ATGACAAAAC GTTCCCAACG TGAGTAGTCA TACCAAGTTC TTTGGCGATA	1020
	TGGTAGGCTT TATCAAGCAA ATCAAAGCTA GCAATTTGTG GAAAATCGTA CTGTGGCCAG	1080
50	TCATTACGAA CGATGTTTGA GTTGGTTGCA GCCGCTGCG CCAAACTAA TTCACGAACA	1140
	TGAACCTCTT CATTCAAAGA ACCTGCAGTT CCCACACGAA TCAATTTCTT CACACCGTAG	1200
55	TCTACGATTA ACTCAGCGC ATAAATCGAA ATAGATGGCA TTCCCATCCC AGTTCCCATG	1260
	ACAGATACAC GGTGACCCTT GTAAGTACCA GTGTAACCAA ACATGTTACG CACTTCGTTA	1320
	AAACAAACAG CATCACCAG GAAATCTCC GCAATAAACT TAGCACGAAG AGGATCCCCA	1380
60	GGAAGAAGAA TTTTATCAGC AATTTACCCC TGCTGAGCAG CAATATGGAT AGACATAATT	1440

-119-

5 TATGATACAA AGAGCGAGAA GAAAACGACT GAAAATTAGG AACCTGACGA GAAATCCTGA 1500  
TTTTTCAGTC AGATTATCTA TTTTCCGAGT TTTCCGCTCG GTTCAAATC AAAACACACG 1560  
CTCTACCTTT CTTTATTTTA TATTTTATAT TGAGAAAGAT ACCAAACCCA TCAAAAAGCG 1620  
AAGGGAAAAT AGGAGTTGGG CGCAGTGAGC GATGCTCGCT AGACCAACTA TCTTTTCC 1680  
10 ACTGCTTTTA GGGTGGGGTC AATTCCTTTC TTTCTTAATT TTGATTTAGA GGAGAGTCGC 1740  
CCGTATTCAG TTCAGCGAAT ACAGTTTACC CATCCTTTCG TTTTATTTT TAGAAAAGTT 1800  
TTCTACTCGT GTTCAAATTA GAACACGCGC TCTACCTTTC TGTTTATACT CTTGAAAAT 1860  
15 CTCTTCAAAC CACGTCAACG TCGACTTGA TATATATGT GACTGACTTC GTCATCTTTA 1920  
TCTACAACCT CAAAG 1935

20 (2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2221 base pairs  
(B) TYPE: nucleic acid  
25 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

30 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

40 TATTATTTT CCCATCCTAA CTGGAACCTA TGTCGCGCGT GTCTTGGACC GAACTGACTA 60  
TGTTACTTC AACTCAGTCG AACTATTTT GTCATTTTTC TTGCCCTTG CAACTTATGG 120  
TGTCTATAAC TACGGTTTAA GGGCTATCAG TAATGTCAAG GATAACAAAA AAGATCTTAA 180  
45 CAGAACCTTT TCTAGTCTTT TTTATTTGTG CATCGCTTGT ACGATTTTGA CCACTGCTGT 240  
CTATATCCTA GCCTATCCTC TCTTCTTTAC TGATAATCCA ATCGTCAAAA AGGTCTACCT 300  
TGTTATGGGG ATTCAACTCA TTGCCCAGAT TTTTCAATC GAATGGGTCA ATGAAGCTCT 360  
50 GGAAAATTAC AGTTTCTCT TTTACAAAAC TGCCTTCATC CGTATCCTGA TGCTGGTCTC 420  
TATTTTCTTA TTTGTTAAAA ATGAACACGA TATTGTTGTC TATACACTTG TGATGAGTTT 480  
55 ATCGACGCTG ATTAACCTACC TGATTAGTTA TTTTGGATT AAAAGAGACA TCAAATTGT 540  
TAAAATTCAC CTAAGTGATT TTAACCACT CTTTCTCCCT CTGACAGCCA TGTTAGTCTT 600  
TGCCAATGCC AATATGCTCT TCACTTTTTT AGATCGCCTC TTCCTCGTTA AAACAGGGAT 660  
60 TGATGTCAAC GTTAGTTACT ATACCATAGC TCAGCGAATT GTGACCGTTA TAGCTGGGGT 720

-120-

5      TGTAACAGGT GCAATTGGAG TGAGTGTGCC TCGTCTCAGT TACTATCTGG GGAAAGGAGA      780  
CAAAGAAGCC TATGTTTCTC TGGTTAATAG AGGTAGTCGA ATCTTTAACT TCTTTATCAT      840  
TCCACTGAGT TTTGGACTCA TGGTTTTAGG ACCAAATGCC ATCCTACTTT ACGGTAGTGA      900  
AAAATATATC GGAGGCGGCA TCTTGACCTC TCTCTTCGCT TTTCGTACGA TTATCCTGGC      960  
10    CTTAGATACC ATTCTTGGTT CCCAAATTCT CTTTACCAAT GGCTATGAAA AACGTATCAC      1020  
AGTCTATACA GTCTTTGCTG GGCTACTCAA TTTGGGCTTG AATAGTCTCC TTTTTTCAA      1080  
CCATATCGTG GCTCCTGAAT ACTACTTACT GACAACTATG CTATCAGAGA CTTCTCTACT      1140  
15    TGTTTTCTAT ATCATTTTCA TCCATAGAAA ACAACTCATC CACTTGGGAC ATATCTTTAG      1200  
CTATACTGTT CGATACTCTC TCTTTTCACT TTCCTTTGTA GCAATTTATT TCCTGATTAA      1260  
20    TTTCGTGTAT CCTGTAGATA TGGTCATTAA TTTGCCATTT TTGATTAATA CTGGTTTGAT      1320  
TGTCTTGCTA TCAGCTATCT CTTATATTAG TCTACTTGTC TTCACAAAAG ATAGCATTTT      1380  
CTATGAATTT TTAAACCATG TCCTAGCCTT AAAAAATAAA TTTAAAAAAT CATAGGAGTT      1440  
25    TAAAATGAAA CAACTAACCG TTGAAGATGC CAAACAAATT GAATTAGAAA TTTTGGATTA      1500  
TATTGATACT CTCTGTAAAA AGCACAATAT CAACTATATT ATTAACACG GACTCTGAT      1560  
30    TGGGGCGGTT CGACATGAGG GCTTTATCCC TTGGGACGAC GATATTGATC TGTCCATGCC      1620  
TAGAGAAGAC TACCAACGAT TTATTAACAT TTTTCAAAAG GAAAAAGCA AGTATAAGCT      1680  
CCTATCCTTA GAAACTGATA AGAACTACTT TAACAACCTT ATCAAGATAA CCGACAGTAC      1740  
35    GACTAAAATT ATTGATACTC GAAATACAAA AACCTATGAG TCTGGTATCT TTATCGACCT      1800  
GCAGGCATGC AAGCTTGGCG TAATCATGGT CATAGCTGTT TCCTGTGTGA AATTGTTATC      1860  
40    CGCTCACAAT TCCACACAAC ATACGAGCCG GAAGCATAAA GTGTAAAGCC TGGGGTGCCT      1920  
AATGAGTGAG CTAACTCACA TTAATTGCGT TGCCTCACT GCCGGCTTTC CAGTCGGGAA      1980  
ACCTGTCTGT CCAGCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTCGCGTA      2040  
45    TTGGGCGCCA GGGTGGTTTT TCTTTTTCAC CAGTGAGACG GGCAACAGCT GATTGCCCTT      2100  
CACCGCCTGG CCCTGAGAGA GTTGACAGCA GCGGTCCACG CTGGTTTGCC CCAGCAGGCG      2160  
50    AAAATCCTGT TTGATGGTGG TTCCGCAAAT CGGCAAAATT CCTTATAAAA TCAAAAGGAA      2220  
T      2221

(2) INFORMATION FOR SEQ ID NO:59:

55

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1509 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

60

-121-

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

	TGAATTTTGA ACAGTACACA GAATACTAAA ATATTCTAG AAATTAATTT GAATTTTCTA	60
	ATTGGATTG TCGCATCTTA TTTCAATCTA CTATAGAAAA AAGTCTTTAA AATTATAAAA	120
15	CGCATCATAT CAAGGTTTTT CAAAAACCTT GATATGATGC CTTTATTGT GGGAATATGT	180
	ATTCATTTT CTAATAAAAT TATGTTTTTG AATAACCTCT ATCTTAGTAG TTTGTATAAT	240
20	CCCCCTCAAT CAGTTTTTGC GATAAGCTTT AATGCTATGA CTATACCACT CTTGCATTTT	300
	TTTTGATGGT GGTGTCATAT AATCGCCATA CATCTGGGT AAAAATTGGT CATATTTTTT	360
	GGGAACAGGC AACATACGGC CCTCAAATC AGTTAAATC AGTTCTTTAA AGGTATCAAC	420
25	TGGGAAGATT TCTTTCATCC CTTCTTACC GATCCCACT CCTCCTCAT ATTGAGGAGT	480
	GTTGGTTACA GCATTTTGA CTAGTTGATC AATTTCTTG TAAAAGTAGC GAGGATTGAC	540
30	AAATCGGAGA GCGTACCAGC TACATAATCT AAGAAAATCT TTTAGTTTGC TATCACCGTG	600
	AACTGCTCGT GATTTTTTGA TATAAGCTAG TTGACGAAGA GCCACATACT TATAGCTCTT	660
	GTCGACAATG CTCAAGTCTG TAAATCGATC AATTGGGAAG ACATCGATGA AAAGGCTGGT	720
35	ATCATGACGC TTGTAATTAA CATGGTCTTC TATAACAGTA GAAGTGCCA AAATCGATGC	780
	GAAATTATGG AAGTACCAAG AAGATGTATC GTAGGAAAGA ACCTTGTAGC GAGGGTGATT	840
40	TTCTTCTTCA ATAATCTTCA GTAAACGCTC ATAATCCTCA CGATAAAGGG AAATATCAAT	900
	ATCATCATCC CAAGGAATCA TACCTTTGTG GCGGATGGCT CCAAGCATGG TTCCATAACT	960
	GAGAAAATAA GGAATATCAT GTTCTTACA AGTCTCATCA ATATAGTCCA GCAGGGCTAG	1020
45	TTGAATTTCT TTAATTTCTT TTTTCTTAA ATATTGCATC CTAATCCTCC AATTATAAG	1080
	CGTGAAATTC ATGACTGTAG AAGCGTTTTT CTTCTGGTGG TAGGGTCATA TAATCTCCAT	1140
50	AAAATTGTGT CAAATAGTA TCAAATTTT CAGGTGCAGG AAGGCTTAAA TTCTCAAAGG	1200
	GTAAATCGAT TGTTTTATCA AAGGTACCAC TTGGGAAGAC TTCCTTTTCC TTAAATTTTG	1260
	AAGGGATAAA AGCCATATAT TGCCCATTTT CACGACTATA TTTTGAATT TCTTCTCGA	1320
55	TTTTATTTGC AAAATAACGA GGAGAAACCG GTCGAAGGAG TAACCAGAAG GCTGTTCGTA	1380
	TCCAATCTTT TAAAAGGCTA TCCTTATAGA CAATATTTT ATGTTTACTG AAAGACAGCA	1440
60	GTTTGAAGCT TTCCAGTTTA TAACAAGTAT CAATGACCTT AGGATCATCA AAGCGATCTA	1500

-122-

TAGGGAAAA

1509

## (2) INFORMATION FOR SEQ ID NO:60:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 671 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

ACAAGGGATT TATTCCTTGG GACGACGACC TAGACTTTTT TATGCCTCGT AAAGATTATG	60
AGAAATTAGC AGAATTATGG CCTCGTTATG CAGATGAACG TTATTTCTTG TCAAAGAGTC	120
25 ACAAGGATTT TGTTGATCGT AATCTTTTAA TTACCATTCTG TGACAAGAAA ACCACCTGTA	180
TCAAGCCTTA TCAGCAGGAT TTGGATTGTC CACATGGTCT GGCCTTGGAT GTTTTGCCTT	240
30 TGGATTATTA TCCGAAAAAT CCAGCTGAGC GGAAAAAACA GGTTCGTTGG GCCTTGATTT	300
ATTCACTCTT TTGTGCGCAA ACTATTCCAG AAAAGCATGG TGATCTCATG AAATGGGGAA	360
GTCGCATTTT ACTGGGTTTG ACTCCAAAAT CTCTCCGTTA TCGCATCTGG AAAAAAGCTG	420
35 AGAAAGAAAT GACTAAGTAT GATTTGGCTG ATTGTGATGG CATTACAGAA TTATGCTCAG	480
GTCCTGGCTA CATGAGAAAC AAGTACCCAA TCACATCTTT TGAAGACAAT CTTTCTTGTC	540
40 CATTTGAAGG AACAGAGATG CCTATTCCAA TCGGCTATGA TGTCTATCTC AGAACTGCTT	600
TTGGGGATTA TATGACGCCT CCACCAGCAG ACAAGCAGGT ACCGCATCAT GATACTGTCA	660
CTGCTGATAT G	671

45

## (2) INFORMATION FOR SEQ ID NO:61:

- 50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2766 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

55

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

60

-123-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

5	ATCTTATACA AGTCGTAAGC CGCTTCCTTA AAACCAGCTT CTAGTAATTC TTCCAATAAG	60
	ATAGTAACCT TCACACCATT TGGTGTTCCC AGTGAATAAA GCTGAAAAGC TTGTTCCTCT	120
	TTTGGCAAGT TTTGTTTGAA ACGGGCACCT GCTGTTGGTC TGTTTAGCCC CGTAAAAGCT	180
10	CCTTGATTAC TAGCTTCATC CTGCCATACG GTCGGTAATT GATATGCTGA CATCCGAGAC	240
	CTCCCTTAAA TCGCATTCTT GTCAAAACCG AGTTTGCGTT GAATAAACTT AACGATTTCTG	300
	ACGATGATAA TCATTGAGAA GCTTCCAGCC ATAACAATTC CCCATTGTGA CAAGTCTAGT	360
15	TTGGTTACGT GGAAGATTCC TTCAAGCGGT TCTACAACGA TTGTTGCCAT GAGAAGGATA	420
	AAGGATACCA AGATGGACCA GTTAAAGGTC TTAGACTTGA ATGGGCCAAC TGTCAAGATG	480
20	GATTGGTAGA CAGACTTGAC ATTGTAGGCA TGGAAGAGCT GAATCAAACC AAGGGTTGCA	540
	AAGGCCATCG TTAGGGCATC TGCATGAATA GCATGATTGT CACCCACATG AACTGGGTAA	600
	GCAATCGCAA GGCCATAAAC ACTCATAACA AGAGCTGCTT GGAGTACACC TTGATAAATG	660
25	ATAGAACTCA AAACACCACC TGAGAAGAAG CTTGCCTTGC GTCCACGTGG TTTATGATTC	720
	ATGACACCAG GTTCCGCAGG TTCAACACCA AGAGCGATAG CTGGGAAGGT ATCCGTTACC	780
30	AAGTTGATCC ACAAAGATG AACCGGCTGC AAGACATCCC AACCAAACAA GGTGATAGG	840
	AAGATGGTTA ATACTTCAGC AGTATTAGCA GAAAGTAGGT ACTGAATAGT CTTTTGAATG	900
	TTTGAGAAGA CCTTACGTCC TTCTTCCACT GCGACGATAA TAGTCGCAA GTTATCATCT	960
35	GCAAGAATCA TATCAGAAGC CCCCTTAGAA ACCTCTGTAC CAGTGATTCC CATACCGATA	1020
	CCAATATCGG CTGTTTTTCA AGCTGGCGCG TCATTGACAC CGTCACCTGT CATGGCAACG	1080
40	ACTTTACCTT GTTTTTTGCA AGCCTTGACG ATACGAACCT TGTGTTCTGG AGACACACGG	1140
	GCATAAACAG AGTATTGACC AACAACTTTT TCAAATTCTT CATCTGACAG TTCATTGAGT	1200
	TCAGCACCAG TTAAACGTG ACCTTCTGTA TCGTTTTCGT CAATGATTCC CAAACGTTTG	1260
45	GCAATGGCTT CCGCTGTGTC TTGGTGGTCA CCTGTAATCA TAATTGGACG GATTCCCGCT	1320
	TCCTTAGCCA CACGAACAGC CTCAGCGGCT TCAGGACGTT CAGGGTCAAT CATCCCAATC	1380
50	AAACCAGTAA AAATTAAATC ATTTTCAAGC TCTTCAGAAG TGAGATTTTC TGGAATACTA	1440
	TCGATAATCT TATAAGCACC TGCAAGGACA CGCAAGGCTT GATGAGCCAT TTCAGAATTG	1500
	TTTGTATGAA TGAGATTTGT AACCTTCTCA TCAATCGGAG CAATATCCCC AGCCTTATCA	1560
55	CGAAGAAGAC AACGTTTTAA GAGTTGGTCT GCGCACCCCT TGACTGCTAC AAGGAAACGA	1620
	CTATCTGGCA ATGGGTGAAC TGTGACATG AGCTTACGGT CAGAGTCAAA TGGCAATTCA	1680
60	GCTACACGAG GATATTTCTC TAAGAAACCT TTGACATCAT AGCCCTTGTC CAAGGCATAT	1740



-124-

TGGATAAAGG CTGTTTCGGT TGGGTCACCA ATCAAGTTAC CTTCCACATC GATTTTCGTA 1800  
TCATTGGCCA AGACAACTGA ACGAAGTAGT GGCATTTCAA GACCTAGTTC AATATCATCA 1860  
5 GCTGAGTCAT GTAGAACCGC ATCGTAGAAG ACTTTTTCGA CTGTCATCTT GTTCATAGTC 1920  
AGCGTACCAG TCTTATCAGA AGCGATGATT TCAGTTGAAC CAAGTGTTTC AACTGCTGGC 1980  
AACTTACGAA CGATGGAATG TCGTTTGGCC AAAACTTGAG TACCAAGAGA AAGAACGATG 2040  
10 GTAACGATAG CAGGAAGTCC TTCTGGAATG GCTGCAACAG CAAGGGCAAC AGAAGTCAAC 2100  
AACTCACCAA GTGGATTTTT CCCTTGAATG AAGACACCCA CTACAAAAGT AACAAGGGCA 2160  
15 ATGACCAAGA TAGCATAGGT CAAGACCTTA GAAAGGTTGT TCAAGTTTTG TTTGAGTGGT 2220  
GTATCAGTCT CATCCGCATC TTGAAGCATA CCAGCAATAT GACCAACTTC AGTATACATA 2280  
CCTGTATTGA CAACAACACC CATCCCACGA CCATAGGTTA CGTTTGAGTT TTGGAAGGCC 2340  
20 ATGTTGACAC GGTCAACCAAT GCCAGCATCT GTCGCAAGAT CGACTGACAA GTCTTTTTTCG 2400  
ACTGGTACAG ATTCACCTGT CAAGGCTGCT TCTTCAATTT TAAGAGAGTT GGCTTCTATC 2460  
25 AAACGTAGGT CCGCTGGTAC CACGTCACCT GCTTCAAGGG CAACGATATC GCCTGGTACC 2520  
AATTCTTTAG AGTCAATCTC TGCCATGTGT CCATCACGAA GAACGCGGGC AACTGGACTA 2580  
GACATGGATT TGAGGGCTTC AATAGCTTCT TCAGCTTTTC CTTCTTGGA AACACCAAAG 2640  
30 GCAGCGTTGA TGATAACCAC AGCTAGGATG ATAATGGCAT CTGCGATATC TTCCCCACCA 2700  
GAAGTCACGA CTGACAAGAT TCTGCCGCAA CTAGGATGAT AATCATCAAA TCCTTAAATT 2760  
35 GCTCGA 2766

## (2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:  
40 (A) LENGTH: 1577 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

TGGATTTATC CTCTTTTTCG TTCTTTTGGG AGCAGTTTTT GAGGAAAAAA TGAGAAAAAA 60  
TACGTCCCAA GCTGTGGAGA AATTACTGGA CTTGCAAGCT AAAACCGCAG AAGTCTTGAG 120  
60 TGATGATAGT TATGTCCAAG TTCCTTTGGA ACAAGTCAAG GTAGGCGACC TGATTCGAGT 180

-125-

	GCGTCCCGGT	GAAAAGATTG	CTGTTGATGG	TGTCGTAGTA	GAAGGTGTCT	CTAGTATTGA	240
	CGAATCCATG	GTGACAGGTG	AGAGTCTGCC	TGTGGACAAG	ACAGTTGGAG	ATACTGTCAT	300
5	TGGCTCAACC	ATCAATCATA	GTGGAACGCT	TGTCTTTAGA	GCAGAAAAAG	TTGGCTCAGA	360
	GACTGTTTTG	GCTCAGATTG	TGGATTTTGT	GAAGAAAGCT	CAGACAAGTC	GTGCGCCGAT	420
	TCAGGACTTG	ACGGATAAGA	TTTCAGGGAT	TTTTGTCCCA	GTAGTTGTCA	TTTtagGAAT	480
10	CATGACCTTT	TGGGTTTGGT	TCGTCTTGCT	CAGGGATAGT	GTGGTCGTGC	TTGGAGCTAG	540
	CTTTGTGTCC	TCTCTTCTCT	ACGGAGTGGC	GGTTTGATTA	TCGCCTGTCC	TTGTGCCTTG	600
15	GGACTTGCAA	CACCGACAGC	CCTTATGGTG	GGGACAGGAC	GTAGTGCCAA	GATGGGGGTT	660
	CTCCTCAAAA	ATGGAAGTGT	CTTACAGGAA	ATCCAGAAAG	TTCAAAGTCT	TGTCCTTGAT	720
	AAGACCGGGA	CTTTGACGGA	AGGGAAACCT	GTGGTAACAG	ATATCATCGG	CGACGAAGTA	780
20	GAAGTGTTTG	GATTGGCAGC	CTCCTTGGA	GATGCTTCTC	AACACCCACT	GGCTGAGGCT	840
	ATCGTTAAGC	GAGCGAGTGA	AGCTGGACTT	GAGTTTCAAA	CTGTTGAAAA	TTTTCAGGCC	900
25	TTGCACGGGA	AAGGTGTTTC	AGGGCGAATC	AATGGAAAAC	AAGTTTTTACT	TGGAATGCT	960
	AAAATGCTGG	ATGGCATGGA	TATTTCTAAT	ACTTATCAAG	ATAAACTAGA	AGAACTAGAA	1020
	AAAGAAGCTA	AGACAGTTGT	GTTTTTAGCT	GTTGACAATG	AAATCAAAGG	CTTGCTTGCT	1080
30	TTGCAAGATA	TTCCTAAGGA	AAATGCTAAG	CTAGCCATCA	GTCAGCTAAA	AAAACGTGGT	1140
	CTCCGAACAG	TCATGCTGAC	AGGAGACAAT	GCTGGTGTGG	CGCGTGCTAT	TGCAGATCAA	1200
35	ATCGGAATTG	AAGAGGTCAT	TGCAGGCGTC	TTGCCAGAAG	AAAAAGCCCA	TGAAATCCAT	1260
	AAACTGCAAG	CGGCTGGCAA	AGTAGCCTTT	GTTGGGGACG	GTATCAATGA	CGCTCCTGCC	1320
	CTTAGTGTAG	CAGATGTGGG	AATTGCTATG	GGTGCTGGAA	CAGATATCGC	CATCGAGTCA	1380
40	GCAGATTTGG	TGTTGACAAC	CAATAATCTT	TTAGGAGTGG	TTCGTGCCTT	TGATATGAGT	1440
	AAGAAAACCT	TTCATCGAAT	TCTACTCAAT	CTTTTCTGGG	CTTTTATCTA	CAATGTTGTC	1500
45	GGAATCCGA	TTGCAGCAGG	AGTCTTTTCA	GGTGTGGCT	GGCTCTCAAC	CCAGATTGGC	1560
	AAGGCTAGCC	CAATGGC					1577

(2) INFORMATION FOR SEQ ID NO:63:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1089 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

55

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

60

(iv) ANTI-SENSE: NO

-126-

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

AAAATGATAT AATAGAATTT ATGGATAAAA ATAAGATTAT GGGATTAACC CAAAGAGAAG 60  
TCAAGGAAAG ACAGGCTGAG GGTTCGGTCA ATGACTTTAC CGCATCAGCC AGTACCAGCA 120  
10 CTTGGCAAAT CGTTAAACGA AATGTCTTTA CCCTTTTAA CGCTTTGAAC TTTGCCATTG 180  
CTTTGGCCCT TGCCTTTGTG CAGGCTTGGA GCAATCTGGT CTTCTTTGCT GTTATCTGCT 240  
15 TTAACGCTTT TTCTGGGATT GTGACCGAGC TACGAGCCAA ACACATGGTG GACAAGCTCA 300  
ATCTCATGAC CAAGGAAAAG GTCAAAACCA TCCGTGATGT CAGGAAGTTG CTCTTAATCC 360  
TGAAGAATTA GTGCTAGGAG ATGTCATTCTG TTTGTCTGCA GGAGAGCAGA TTCCTAGTGA 420  
20 TGCCTTGCTT TTGAAGGCT TTGCGGAAGT CAATGAAGCC ATGTTAACGG GAGAAAGTGA 480  
TTTGGTGCAA AAGGAAGTTG ACGGCTTACT TTTGTCAGGA AGTTTCCTAG CCAGTGGGTC 540  
25 AGTTTTATCT CAAGTTCACC ATGTCGGTGC AGACAACAT GCTGCCAAAC TCATGCTTGA 600  
GGCTAAGACC GTTAAACCCA TCAACTCCCG TATCATGAAA TCGCTGGACA AGTTGGCTGG 660  
TTTTACTGGG AAGATTATCA TTCCCTTTGG TCTGGCTCTC TTGCTGGAAG CCTTGCTTTT 720  
30 AAAAGGCCTG CCTCTCAAGT CATCCGTTGT AAACCTCGTC ACAGCTCTTT TGGGAATGTT 780  
GCCTAAGGGA ATTGCCCTTT TGACCATTAC TTCGCTCTTG ACTGCAGTGA TTAAGTTGGG 840  
35 CTTGAAAAAG GTCTTGGTGC AGGAGATGTA CTCTGTTGAG ACCTTGCGC GCGTGGATAT 900  
GCTCTGTCTG GACAAGACGG GCACCATCAC CCAAGGAAAG ATGCAGGTGG AGGCTGTTCT 960  
TCCACTGACG GAAACTTACG GTGAAGAGGC TATTGCCAGC ATCTTGACTA GCTACATGGC 1020  
40 CCATAGTGAG GATAAGAATC CAACTGCCCC AGCCATTCGC CAGCGTTTGT GGGAGATGTT 1080  
GCTTATCCT 1089

## 45 (2) INFORMATION FOR SEQ ID NO:64:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 731 base pairs  
(B) TYPE: nucleic acid  
50 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## 55 (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

60

-127-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GCTAGCAATA TCATGTTTAT GCTTGATTG GGAATCATT TAGATCAGTG GTCCTTGAAA 60  
5 AAAACTGCAA CAGATTTAGA ACAGAGTCTT CTTGCAAAAG AGAGCGATGT ATTCCTAGTA 120  
CAGGGCGATA CGGTTGTTAG TATCAAGAGT TCCGATGTTC AAATAGGAGA TGTCTTGATC 180  
10 TTATCTCAAG GAAATGAAAT TCTGTTTGAT GGACAAGTAG TTTCAGGTTT AGGTATGGTC 240  
AACGAAAGTT CCTTGACAGG AGAGAGTTTT CCAGTTGAAA AAAGAGAGTC TGATTTGGTT 300  
TGTGCAAATA CAGTATTAGA AACTGGAGAG TTACGCATTC GTGTAACAGA TAATCAGATG 360  
15 AACAGCCGTA TTTTACAGCT GATTGAGTTG ATGAAGAAAT CTGAAGAAAA CAAGAAAACG 420  
AAACAACGCT ATTCATCAA GATGGCGGAT AAAGTCGTCA AATATAATTT CTTGGGGTCT 480  
GGGCTGACTT ACCTATTGAC AGGTTCTTTT TCTAAGGCTA TTTCTTTCCT ATTGGTCGAT 540  
20 TTCTCCTGCG CTTTGAAAT CTCTACTCCT GTAGCTTATT TGACAGTTAT CAAGGTAGGG 600  
TTGAACCGTG AAATGGTGAT TAAGGATGGA GATGTTCTGG AGAAATATCT GGTAGTTGAT 660  
25 ACTTTCTTGT TTGATAAGAC AGGACCAATC ACAACTAGTT ATCCTATAGT TGAAAAGGTG 720  
TACCCTTTGG G 731

## (2) INFORMATION FOR SEQ ID NO:65:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2197 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

40 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

TATATTATTC CATTTGTGGT AAATCTGTAC ATGATAGATT AAGTACTCCG ACTGAAACCA 60  
50 GTACACTAAT CAAGCTATAG CCAGCTAACA AAAGGAGTAA CCATAGAATA TTAACTTTAA 120  
AATTTTCCTT CATCGTTTAC ACCTTCTCTT TCACATTCTT ACCAAGGATA CCAGCTGGGC 180  
55 GGACAATCAA GATCAACAAC AAGATTCCAT AAACAATGGC ATCACGGAAA TCTGACATCC 240  
CAAAGGCTGT CGCAAAGGTT TCCAATAGAC CAATCACAAA GCCACCAAGA GCCGCACCAG 300  
GAATAATTCC GATACCACCA AGTACTGCGG CAACGAAAGA TTTAAGACCT GGAGTAACCC 360  
60 CCATCAAAGG CTCAAGAGAG TTATAATAAA GAGCAATCAG AACACCAGCC GCACCCGCAA 420

-128-

	GAGCAGAACC CAAAGCGAAG GTAAAGCTAA TCGTACGGTT TATATTGATC CCCATCAATT	480
	GCGCCGCGTC GCTATCTACT GATACTGCAC GCATGGCTTT CCCCATCTTA GTCTTTTGA	540
5	CAATGACTTG TAACAAAATC ATCAAAATCA AGGAAATGCC CAAAATCATT AACTGCACAT	600
	TTGTTAAGCT AATTGGTCCC AAATCATATC GAACTGTTTG AATCGCTTGA GGGAAGGCAC	660
	GGGTATTGGC ACCAACCAGA TAGACCATT CATACTCCAA TAGGAAAGAA ACCCCAATAG	720
10	CCGTAATCAA AACAGCAATA CGAGTAGAGT GGCGCAAAG TCGGTAAGCA AGAAACTCAA	780
	TCACGACACC AAGAATAGCT GTCGCTAGCA TAGCTACAAT AAGCGCTACA AAGAAATTCA	840
15	TTTGGAAGA ATTGATCAAG AAATAACCGA TAAAGGCTCC CATCATATAA ATATCACCAT	900
	GGGCGAAGTT GATGAGCTTG ATAATTCCGT AAACCATGGT ATATCCTAGG GCTAACAGCG	960
	CGTAAACACT ACCTAGAATC AAACCATTTA CGAGTTGTTG GAGCATAAGA TTCACTCTTT	1020
20	CTATTTATAA TTCCGAGGGT TTTCCCTCAC TTTTGTAGT GTTCTTATAC TCAATGAAAA	1080
	TCAAAGAGCA AACTAGGAAA CTAGCCGCAG GTTGCTCAA GCACTGCTTT GAGGTTGTAG	1140
25	ATAAGACTGA CGAAGTCAGT CACATATATA ATCCAAGGCG ACGTTGACGC AGTTTGAAGA	1200
	GATTTTCGAA GAGTATTAAA TATCGAAACA GGGAGTGAGT CAAAGGCTCA TTCCCTATTT	1260
	CAACATTTTT CTATTATGGT TTTACAACCT CTGCTGCTTC AACTTTACCA TTGTTTATGG	1320
30	TCATCATGTA AGCAGTTTTG ACTGTGTTGT GGTCTGCATC GAAGCTTGTT TGACCAGTTA	1380
	CACCTTCAA ATCTTTTGTT TTAGCAAGGT TATTCTTGAT TTCACCTGAA TTTTTCAGC	1440
35	CTTTTGCTGC GTTTGCTACA AGGTGAACTG AATCATAAGC CAAGGCTGCA AATGTTGAAG	1500
	GCTCTTCATT GTACTTAGCA CGGTAAGCGT CAAGGAAGGC TTTAGCTTTA GCTGAACTT	1560
	CTACAGTAGT TGAGAAGCCT GAGATAAAGT AGATGTTTGA TGCTTTTTC GAGTTGCTT	1620
40	GTTGTACAAA CTCCTACCG TTGAATCCAT CACCACCAAC GATTGGTTTG TCAATTCCCA	1680
	TACCACGCGC TTGGTTTACA ATCTTACCAG CCTCATTATA GTAACCAGGA ACAACGATAG	1740
45	CATCAAAGTC TTTCCCTTTC ATTTTGTAA GGGCTGCTTG GAAGTCTGTG TCACCTGCTA	1800
	CGAAAGTTTC ATCTGCAACG ATTTACCCCT TGTATGACTC GCGGAAAGAT TTGGCAATCC	1860
50	CTTTAGCATA GTCAGTGGCA TTGTCAAGT AAAGAACAAC TTTCTTAGCA TTTAATTTTT	1920
	CAGAAACATA GTTTGAGATA ATTTTTCCTT GGAAGCTATC TTGGAAAGTT CCAATAAAGA	1980
	GGTAATCTTG ACCTTTAGTC AATCCATCTT GAGTCGCACT TGGTGAGATC AATGGAACAC	2040
55	CTGCTTTTGT AGCGTTCGCT ACCGCAGCTG CAGTCGCACC AGATGTCGCA GGTCTACGA	2100
	CTGCTGATAC TTTAGATTGG GTTACAAGGT TAGTTGTAAC TGAAGCAGCC TCAGCTGTTT	2160
60	CAGACTTATT ATCTTTATCG ACTACTTCGA TTTGTTT	2197

(2) INFORMATION FOR SEQ ID NO:66:

-129-

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 900 base pairs  
    (B) TYPE: nucleic acid  
5      (C) STRANDEDNESS: single  
        (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 10    (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 15
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:
- |    |  |     |
|----|--|-----|
| 20 | TGTCCCAAGA CCAGACTTGG TATGCTCTGG CCTATGATGG GGCAGAAGTG ATTGGCTTTC  | 60  |
|    | TAAGTGTTCG GGAGACTCTC TTTGAAGCAG AAGTCCTGCA AATCGCTGTC AAAGGAGCTT  | 120 |
|    | ATCAGGGTCA GGGAAATTGCG TCAGCCTTGT TTGCTCAATT GCCGACAGAC AAGGAAATTT | 180 |
| 25 | TCCTCGAAGT CAGACAGTCA AATCAACGAG CGCAAGCATT TTACAAGAAA GAAAAGATGG  | 240 |
|    | CAGTTATCGC TGAGCGAAAG GCCTACTACC ATGACCCAGT CGAGGACGCC ATTATCATGA  | 300 |
|    | AGAGAGAAAT AGATGAAGGA TAGATATATT TTAGCATTTG AGACATCCTG TGATGAGACC  | 360 |
| 30 | AGTGTCGCCG TCTTGAAAAA CGACGATGAG CTCTTGTCCTA ATGTCATTGC TAGTCAAATT | 420 |
|    | GAGAGTCACA AACGTTTTGG TGGCGTAGTG CCCGAAGTAG CCAGTCGTCA CCATGTCGAG  | 480 |
| 35 | GTCATTACAG CCTGTATCGA GGAGGCATTG GCAGAAGCAG GGATTACCGA AGAGGACGTG  | 540 |
|    | ACAGCTGTTG CGGTTACCTA CGGACCAGGC TTGGTCGGAG CCTTGCTAGT TGGTTTGTCA  | 600 |
|    | GCTGCCAAGG CCTTTGCTTG GGCTCACGGA CTTCCACTGA TTCCTGTAA TCACATGGCT   | 660 |
| 40 | GGGCACCTCA TGGCAGCTCA GAGTGTGGAG CCTTTGGAG TTTCCCTTGC TAGCCCTTTT   | 720 |
|    | AGTTCAGTGG GTGGGGCACA CAGAGTTGGT CTATGTTTCT GAGGCTGGCG ATTACAAGAA  | 780 |
| 45 | TTGTTGGGGA AGACACGAGA CGATGCAGTT GGGGAGGCTT ATGACAAGGT CGGTCGTGTC  | 840 |
|    | ATGGCTTGAC CTATCCTGCA GGTCGTGAGA TTGACGAGCT GGCTCATCAG GGCAGGATA   | 900 |
- 50    (2) INFORMATION FOR SEQ ID NO:67:
- (i) SEQUENCE CHARACTERISTICS:  
            (A) LENGTH: 1023 base pairs  
            (B) TYPE: nucleic acid  
55          (C) STRANDEDNESS: single  
            (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 60    (iii) HYPOTHETICAL: NO

-130-

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

	CCGGCGATCT TCCGCTAGAA ATAGTCTACC AAGATGAGGA TGTGGCTGTC GTTAACAAAC	60
10	CTCAGGGAAA TGGTTGTGCA CCCGAGTGCT GGTACATACCA GTGGAACCCT AGTAAATGCC	120
	CTCATGTATC ATATTAAGGA CTTGTCGGGT ATCAATGGGG TTCTGCGTCC AGGGATTGTT	180
	CACCGTATTG ATAAGGATAC GTCAGGTCTT CTCATGATTG CTAAAAACGA TGATGCGCAT	240
15	CTAGTACTTG CCCAAGAACT CAAAGATAAA AAGTCTCTCC GCAAATATTG GGCGATTGTT	300
	CATGGAAATC TGCCTAATGA TCGTGGTGTA ATTGAAGCGC CGATTGGCCG GAGTGAAAAA	360
20	GACCGTAAGA AACAGGCTGT AACTGCTAAA GGGAGCCTG CAGTGACGCG TTTTCACGTC	420
	TTGGAACGCT TTGGCGATTA TAGCTTAGTA GAGTTGCAAC TGGAGACAGG GCGCACTCAT	480
	CAAATCCGTG TCCACATGGC TTATATCGGC CATCCAGTCG CTGGTGATGA GGTCTATGGT	540
25	CCTGCAAGAC TTTGAAAGGA CATGGACAAT TTCTTCATGC CAAGACTTTA GGTTTTACTC	600
	ATCCGAGAAC AGGTAAGACC TTGGAATTTA AAGCAGATAT CCCAGAGATT TTTAAGGAAA	660
30	CCTTGAGAG ATTGAGAAAG TAAGATGAA AAAGAAATTA ACTAGTTTAG CACTTGTAGG	720
	CGCTTTTTTA GGTGTGTCAT GGTATGGGAA TGTTTCAGGCT CAAGAAAGTT CCAGGAAATA	780
	AAATCCACTT TATCAATGTT CAAGAAGGTG GCAGTGATGC GATTATTCTT GAAAGCAATG	840
35	GACATTTTGC CATGGTGGAT ACAGGAGAAG ATTATGATTT CCCAGATGGA AGTGATTCTC	900
	GTTATCCATG GAGAGAAGGA ATTGAAACGT CTTATAAGCA TGTTCTAACA GACCGTGTCT	960
40	TTCGTCGTTT GAAGGAATTG AGTGTCCAAA AACTTGATTT TATTTTGGTG ACCCATACCC	1020
	ACA	1023

45

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 675 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

55

(iv) ANTI-SENSE: NO

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

-131-

5 GCTCGGTACC CGGGGATCCT CTAGAGTCGA TAATATCAAC CTGCAGGTG ATGAACGAGA 60  
TCGGATTGCT CTTGTGGGA AAAATGGTGC AGGTAAGTCT ACTCTTTGA AGATTTTAGT 120  
TGGAGAAGAG GAGCCAACTA GCGGAGAAAT CAATAAGAAA AAAGATATTT CTCTGTCTTA 180  
CCTAGCCCAA GATAGCCGTT TTGAGTCTGA AAATACCATC TACGATGAAA TGCTTCATGT 240  
10 CTTTAATGAT TTGCGTCGGA CGGAGAGACA ACTGCGTCAG ATGGAGCTGG AGATGGGTGA 300  
AAAGTCTGGT GAGGATTTGG ATAACTGAT GTCAGATTAT GACCGCTTAT CTGAGAATTT 360  
TCGCCAAGCA GGTGGCTTTA CCTATGAAGC TGATATTCGA GCGATTTTGA ATGGATTCAA 420  
15 GTTTGACGAG TCTATGTGGC AGATGAAAAT TGCTGAGCTT TCTGGTGGTC AAAATACTCG 480  
TTTGCACTT GCCAAAATGC TCCTTGAAA GCCAATCTC TTGGTCTTGG ACGAGCCAAC 540  
20 TAACCACTTG GATATTGAAA CCATCGCCTG GCTAGAGAAT TACTTGGTAA ACTATAGCGG 600  
TGCCCTCATT ATCGTCAGCC ACGACCGTTA TTTCTTGAC AAGGTTGCGA CAATTACGCT 660  
AGATTTGACC AGCAT 675

## (2) INFORMATION FOR SEQ ID NO:69:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 582 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

45 TAGAGTCGAT AGCAATAGAT TGAGTAAGTT GTAGCCTTAC CAGCATCGAT AGAAACAAGG 60  
GCACCACGGT GACGTCCACC AATTTCCCCT GGAATCAATG GCAAGTATTG GTCGAAGGTA 120  
TGGTTCATGA TACCGTAACC AUGAGTCATT GACAAGAACT CAGTTGAGTA TCCAATCAAA 180  
50 CCACGCGCTG GAACAAGGAA GACCAAACGA GTTGACCAT TACCAGTTGA AATCATATCC 240  
AACATTCAC CTTTACGTTT AGAAAGGCTT TGGATAACAG ACCCTTGGTA TTCTTCTGGA 300  
55 GTGTCGATTT GTACACGTTT AAATGGTTCA CATTTAACAC CGTCGATTTT TTTTACGATA 360  
ACTTCTGGAC GAGATACTTG AAGTTCATAG CCCTCACGAC GCATTGTTT GATAAGGATT 420  
GACAAGTGCA ATTCTCCACG TCCTGAAACA GTCCATTTAT CTGCGTGAAT CAGTTGGGTC 480  
60 AACACGAAGG AACGTCTGTT TGCAATTCTG CCTGCAAGCG TTCTTCCACC TTACGAGAAG 540



-132-

TTACCCATTT ACCTTCTTTA CCAGCAAATG GTGAGTTGTT GA

582

(2) INFORMATION FOR SEQ ID NO:70:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1337 base pairs

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

TTGGATTGAA GAACAAAGAT TTGGACTCTA TTGACCTTAT GGT TTGGGGG AAATTTGGAA	60
25 TTTCAAAGTC GCCCAACCCC CTCATTTCTA AAGAATTGGA AGCCGGATGG GACTCTACCA	120
AACACGTTTA ACCCAAAGAA AATTGGCAGG AAGAAATGGA AAAAATTGA TTTTAAAAA	180
ATACTTAAGG AAACCTTAAG CTAGGGAGTG TACCCTAAGT TCAATAAAGT TAAAGAAGAC	240
30 CTTAACTTAA ACTCCTAAAA CTTTTTCAAT AATAATCTCC CTATAAAAAT AAAGTCGCCC	300
AATCAGGCGG CTTAATTTTT TTGAAAAATG GGCTTGGTGC CTGAGAATAA ATAGCTTAGT	360
35 GATAGAAGAA AATGGGGAAA TATGGTATAA TGAAACGATA GATTTTGTAA TAGGAATAAG	420
ATCATGTTTG GATTTTTTAA GAAAGATAAG GCTGTGGAAG TAGAGGTTCC GACACAGGTT	480
CCTGCTCATA TCGGCATCAT CATGGATGGC AATGGCCGTT GGGCTAAAAA ACGTATGCAA	540
40 CCGCGAGTTT TTGGACACAA GCGGGGCATG GAAGCATTGC AAACCGTGAC CAAGGCAGCC	600
AACAACTGG GCGTCAAGGT TATTACGGTC TATGCTTTTT CTACGGAAAA CTGGACCCGT	660
45 CCAGATCAGG AAGTCAAGTT TTCATGAAGT TGCCAGTAGA GTTTTATGAT AATTATGTCC	720
CGGAAGTACA TGCGAATAAT GTTAAGATTC AAATGATTGG GGAGACAGAC CGCCTGCCTA	780
AGCAAACCTT TGAAGCTTTA ACCAAGGCTG AGGAATTGAC TAAGAACAAC ACAGGATTGA	840
50 TTCTTAATTT TGCTCTTAAC TATGGTGGAC GTGCTGAGAT TACACAGGCG CTTAAGTTGA	900
TTTCCCAGGA TGTTTTAGAT GCCAAAATCA ACCCAGGTGA CATCACAGAG GAATTGATTG	960
55 GTAACATCT CTTTACTCAG CATTTGCCTA AGGACTTACG AGACCCAGAC TTGATTATCC	1020
GTACTAGTGG AGAATTACGT TTGAGCAATT TCCTTCCATG GCAGGGAGCC TATAGTGAGC	1080
TTTATTTTAC GGACACCTTA TGGCCTGATT TTGACGAAGC GGCCTTGACG GAAGCTATTC	1140
60 TTGCCTATAA TCCTCGTCAT CGCCGATTTG GAGGAGTTTA GGAGGAAATA TGACCCAGGA	1200

-133-

TTTACAGAAA AGAACCTTGT TTGCAGGGAT TGCCCTGGCT ATTTTCCTAC CAATTTTAAT 1260  
GATTGGGGGC TCTTGCTTCA GATAGCAATC GGAATCGTAG CCATGCTAGC CATGCATGAA 1320  
5 CTTTGAAGA TAAGAGG 1337

## (2) INFORMATION FOR SEQ ID NO:71:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 818 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20 (iv) ANTI-SENSE: NO

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

TCGGTACCCG GGGATCCTCT AGAGTCGATA GTCGCCAAGC AGAAGAAGGG AACACCATTC 60  
GTAGAAGACG TGAGTGGCAG GAATGCCAAC ACCGTTTAC AACCTACGAA CGAGTAGAAG 120  
30 AAAGAACCTT AGTGGTTGTT AAAAAAGATG GCACACGGGA ACAATTCTCC AGAGATAAAA 180  
TCTTTAATGG GATTATCCGC TCAGCCCAGA AACGTCCTGT GTCAAGTGAT GAAATCAACA 240  
35 TGGTGATCCT CTAGAGTCGA ACAGAACTC CGTGGTCGAA ATGAAAATGA AATTCAAAGT 300  
GAGGACATTG GTTCACTCGT CATGGAGGAG TTGGCTGAAT TGGACGAGAT TACCTATGTA 360  
CGTTTTGCTA GTGTCTATCG TAGTTTTAAG GATGTCAGTG AGTTAGAGAG CTTGCTCCAA 420  
40 CAAATCACCC AGTCCTCTAA AAAGAAAAAG GAAAGATAAA TGAAGCCAAT TGACCGTTTT 480  
TCTTATCTAA AGAATAATCG GGTGTCGCAA GATACCTCAT CTCTGGTACA GTGCTACCTC 540  
45 CCGATTATCG GTCAGGAGGC ACTGAGCCTT TATCTTTATA CGATTAGTTT TTGGGATAAT 600  
GGTAGAAAGG AATATCTTTT TTCAAGTATC CTCAATCATC TTAACCTTGG AATGGATAGA 660  
CTGATAAAAT CATTGAAAAT CTATCTGCT TTTAATCTCT TGAATCTCTA TCAAAAGGGG 720  
50 GATGTTTATC AGCTAGCCCT CCATGCTCCT CTATCTAGTC AAGACTTCTT GGGGCATCCT 780  
GTTTATCGCA GACTCTTAGA GAAAAAGATT GGGGACAA 818

## 55 (2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 746 base pairs  
60 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

-134-

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

5	TTACCCGGGG ATCCTCTAGA GTCGATATGC TCTCTGAGGG TCAATTCCTC ATACAGACTA	60
15	GGCGTCTCAG GAATGTAGCC AATCTGCTTG CGGTAGCTAG TCGCATCTCC TTGCAGAGTC	120
	AGGCCATTGA TATTGATGGA GCCACTATAA GGTGCCAACA GACCGATAAT CTCATTGATC	180
20	GTCGTTGATT TCCCAGCACC ATTGAGACCA ATCAAACCGA CCAACTGCCC ACTTTCAACA	240
	GTAAAGGACA CATCTTTCAA AACAGGAACA TGAACATAGC CACCTGTCAG GTTTTTAATT	300
	TCTAACATAT TTTCTCCAAA TCTGGTATAA TGTAGCTATA TTATATCAAA ATTCAGTACA	360
25	GTAGAGGTAG ATTTTATGTC AGATTGCATT TTTTGTAATA TCATCGCAGG GGAAATTCCT	420
	GCTTCGAAAG TATATGAAGA TGAGCAGGTC CTTGCCTTTC TTGATATCTC TCAAGTAACA	480
30	CTAGGACACA CCTTGGTCGT GCCAAAAGAA CACTATCGCA ATCTTTTGGG GATGGATGCT	540
	ACGAGCGCCA CCAACTCTTT GCCCAAGTAC CAAAAGTAGC TCAAAAAGTC ATGAAAGTCA	600
	CTAAGGCTGC TGGTATGAAT ATCATTTCCT ACTGTGAAGA AGTCGCTGGT CAAACAGTTT	660
35	TTCATACTCA CGTTCACCTT GTGCCTCGCT ACAGTGCTGA CGATGACCTC AAGATTGATT	720
	TTATCGCCCA CGAAACAGAC TTTGAC	746

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 767 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

55	GATCCAAGCA GTCCGTGATG TAAGCTTTGA AGTTAATGAA GGAGAAGTTG TTTCCCTTAT	60
60	CGGTGCCAAC GGTGCAGGTA AGACAACTAT TCTTCGCACC TTGTCAGGTT TGGTTCGACC	120

-135-

AAGTTCAGGA AAGATTGAAT TTTTAGGTCA AGAAATCCAA AAAATGCCAG CTCAGAAAAT 180  
TGTGGCAGGT GGTCTTTCAC AAGTTCCAGA AGGACGCCAC GTCTTTCCTG GCTTGACTGT 240  
5 TATGGAATCT CTTGAAATGG GAGCTTCTT AAAGAAAAAT CGTGAAGAAA ATCAAGCTAA 300  
CTTGAAGAAG GTTTTCTCAC GCTTTCCTCG TCTTGAAGAA CGTAAGAACC AAGATGCAGC 360  
TACTCTTTCA GGAGGGGAAC AACAAATGCT TGCCATGGGA CGCGCTCTTA TGTCAACACC 420  
10 AAAACTTCTT CTTTATAGATG AACCATCAAT GGGACTTGCC CCAATCTTCA TCCAAGAGAT 480  
TTTTGATATC ATTCAAGATA TTCAGAAGCA AGGAACAACC GTCCTCTTGA TTGAACAAAA 540  
15 TGCCAATAAA GCACTTGCAA TCTCTGACCG AGGATATGTA CTGGAACAGG GAAATCGTCT 600  
ATCAGGGACA GGGAAAGACT CGCTCATCAG AGGAGTCAGA GCATATCTAG GTGGTAAACA 660  
TCCAGTGGAT TTTTGTCCGC AGTGAGTTCG GGATCATCAT TTAGTTGGGG CTTGTTAGGT 720  
20 TCAGTAAGTC GGTATCAAA TCAGGGTTGT TTGCCGAGT GGGGTCG 767

## (2) INFORMATION FOR SEQ ID NO:74:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 695 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
30 (ii) MOLECULE TYPE: DNA (genomic)  
(iii) HYPOTHETICAL: NO  
35 (iv) ANTI-SENSE: NO  
40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

GAGCTCGGTA CCCGGGGATC CTCTAGAGTC GATAATTCGT TGGTTGGACG AACCTCGAAA 60  
CTGGAGCATG AGATTTCTCT TAGTTCGATC ATATCTTCCA TCGACAAGAA TGTCAATCAA 120  
45 TGATAAGAGT TCCAGTTTAT CTGGAGTTTC CGGGATCATT TCTTCCAAG TGTAGCCCGT 180  
CCAGGACCAA ATGTCCTTGT CTGGCAATTC CTTTCGGATG CGTTTAACTA GAGGCAAGAG 240  
50 AATGCCAGTA TTGAGAAAAG GCTCCCCTCC CAGCAAAGTC AAGCCTTGAA CATAGGGTTG 300  
GGCAAGGTCT GCCATAATCT GCTCTTCTAA TTCTGCTGTA TAGGGAATGC CAGCATTAAG 360  
AGACCAAGTC GCAACATTAT AACATCCCTC GCAGTGAAAC ATACAGCCTG ATACATAGAG 420  
55 AGAGTTGCGC ACGCCTTCGC CGTCCACAAA GTTAAAGGCC TTGTAGTCAA TGATACGACC 480  
TTGACTAAGT TCCTCGCTTT TCCATTCTTG TGGTTTGGGA TTATTCATTC GCTACCTCTA 540  
60 TCCAATAACG CTCGACTCCA TTGCGAGCAT CCTCAAATAT TCCACCATTG GCTAGAATGA 600

-136-

CTGCTCTGCT AGCAGGATTA TTCACGCTAC AGGGCACCAG AGCTTTCTTG ATGTCTTTTC 660  
CCTAGCAACT TCAAGCCCTG ACGGAAGTCT TTTT 695

## 5 (2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 723 base pairs  
    (B) TYPE: nucleic acid  
10      (C) STRANDEDNESS: single  
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

CTCGGTACCC GGGGATCCTC TAGAGTCGAC GGCTACAATG ATATTAAGAT GGATGATGTG 60  
25 ATTGACGCGT ATGTCATGGA AGAAATCAAG AGATAAGATT TTTTGCTCCT TTCTTAGGTG 120  
GTGAGGGACG CAAGCAAACC GATGGTTTCA TTGCTTATTT TTGAGCCTAG GGTCTCAAAA 180  
30 ATCCCCTGTG ATGGGACTGA TAAATCAGTT CCATCACTTT CACCACGGCG AAAGAAGCAG 240  
ATGACTTCAA ATTGAACTTC GTTTCATTTT AAAGTAAAA TCAAGAAGTT TAAAATAGCT 300  
AGGTCTGCTG GCCTAGCTTT TGTTCAAAG TAGAGAAAGG AATATCATGG TAAATCATTT 360  
35 CCGTATAGAT CGTGTGGGCA TGGAAATCAA GCGTGAAGTC AATGAGATTT TGCAAAAGAA 420  
AGTCCGTGAT CCACGTGTCC AAGGTGTGAC CATCACAGAT GTTCAGATGC TGGGTGACTT 480  
40 GTCTGTTGCC AAGGTTTATT ACACCATTTT GAGTAACCTT GCTTCGGATA ACCAAAAAGC 540  
CCAAATCGGG CTTGAAAAAG CAACTGGTAC CATCAAACGT GAACTTGGTC GCAATTTGAA 600  
ATTGTACAAA TCCCAGATTT GACCTTCGTC AAAGACGAGT CCATCGAGAT GGAACCAAGA 660  
45 TTGACGAGAT GCTACGAAAT CTGGATAAGA CTAAAGAAGA GGGGGTTGCC CCCCTTTTTT 720  
GGG 723

## 50 (2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 970 base pairs  
55      (B) TYPE: nucleic acid  
        (C) STRANDEDNESS: single  
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

60 (iii) HYPOTHETICAL: NO

-137-

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

	TGTCCTTATT TGTCTGACCA AGTGCAAGCT GGTCCGATTT GTGGTAACAT TGGATAAGAT	60
10	TTGACAAAGG AATTTCCATC ATGTAACGGT CTTACTCCAC GAAACGATTG ATATGCTTGA	120
	CGTAAAGCCT GAAGGTATCT ACGTTGATGC GACTTTGGGC GGAGCAGGAC ATAGCGAGTA	180
	TTTATTAAGT AAATTAAGTG AAAAAGGCCA TCTCTATGCC TTTGACCAGG ATCAGAATGC	240
15	CATTGACAAT GCGCAAAAAC GCTTGGCACC TTACATTGAG AAGGGAATGG TGACCTTTAT	300
	CAAGGATAAC TTCCGTCATT TACAGGCACG TTTGCGCGAA GCTGGTGTTT AGGAAATTGA	360
20	TGGAATTTGT TATGACTTGG GAGTGTCTAG TCCTCAATTG GACCAGCGTG AGCGTGGTTT	420
	TTCTTATAAA AAGGATGCGC CACTGGACAT GCGGATGAAT CAGGATGCTA GTCTGACAGC	480
	CTATGAAGTG GTTAATCATT ATGACTATCA TGATTTGGTT CGTATTTTCT TCAAATACGG	540
25	TGAGGATAAA TTCTCTAAAC AGATTGCGCG TAAGATTGAG CAAGCGCGTG AAGTGAAGCC	600
	GATTGAGACA ACGACTGAGT TAGCAGAGAT TATCAAGTTG GTCAAACCTG CCAAGGAACT	660
30	CAAGAAGAAG GGTATCCTG CTAAGCAGAT TTTCCAGGCT ATTCGAATTG AAGTCAATGA	720
	TGAACTGGGA GCGGCAGATG AGTCCATCCA GCAGGCTATG GATATGTTGG CTCTGGATGG	780
	TAGAATTTCA GTGATTACCT TTCATTCTT AGAAGACCGC TTGACCAAGC AATTGTTCAA	840
35	GGAGCTTCAA CAGTTGAAGT TCCAAAAGGC TTGCTTTCAT CCCAGATGAT CTCAAGCCCA	900
	AGATGGAATT GGTGTCCCGT AAGCCAATCT TGCCAAGTGC GGAAGAGTTA GAAGCCAATA	960
40	ACCGTTGACT	970

(2) INFORMATION FOR SEQ ID NO:77:

45	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 954 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
	(D) TOPOLOGY: linear

50	(ii) MOLECULE TYPE: DNA (genomic)
----	-----------------------------------

(iii) HYPOTHETICAL: NO

55	(iv) ANTI-SENSE: NO
----	---------------------

60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:
----	--

	GAAAGGAGTA ACTGATGCAC GTAACAGTAG GTGAATTAAT TGGTAATTTT ATTTTAATCA	60
--	---	----

-138-

CTGGCTCTTT TATTCTTTTG CTAGTCTTGA TTAAAAAATT TGCATGGTCT AATATTACAG 120  
GCATTTTTCGA AGAAAGAGCT GAAAAAATTG CTTGAGATAT TGACAGAGCT GAAGAAGCCC 180  
5 GTCAAAAAGC AGAAGTATTG GCTCAAAAAC GCGAAGATGA ATTGGCTGGT AGCCGTAAAG 240  
AAGCTAAGAC AATCATTGAA AATGCAAAGG AAACAGCTGA GCAAAGTAAG GCTAATATCT 300  
10 TAGCAGATGC TAAACTAGAA GCGGGGACACT TAAAAGAAAA AGCCAATCAA GAAATTGCTC 360  
AAAATAAAGT AGAAGCTTTA CAGAGTGTTA AGGGTGAGGT CGCAGATTTG ACCATCAGCT 420  
TAGCTGGTAA AATCATCTCA CAAAACCTTG ACAGTCATGC CCATAAAGCA CTCATTGATC 480  
15 AGTATATCGA TCAGCTAGGA GAAGCTTAAT GGACAAGAAA ACAGTAAAGG TAATTGAAAA 540  
ATACAGCATG CCTTTTGTCC AATTGGTACT TGAAAAAGGA GAAGAAGACC GTATCTTTTC 600  
20 AGACTTGACT CAAATCAAGC AAGTTGTTGA AAAAACAGGT CTGCCTTCTT TTTTAAACA 660  
AGTGGCAGTA GACGAGTCGG ATAAGGAAAA AACAATTGCT TTTTCCAAG ATTCTGTGTC 720  
ACCTTTATTA CAAAACCTTA TCCAGGTTCT GGCCTACAAT CACAGAGCAA ATCTTTTTTA 780  
25 TGATGTGCTT GTAGATTGCT TGAACCGACT TGAAAAAGAA ACAAATCGAT TTGAAGTGAC 840  
GATTACGTCT GCTCATCCTC TAACTGATGA ACAGAAGACT CGTTTGCTCC CTTTGATTGA 900  
30 GAAAAAATG TCTCTGAAAG TAAGGAGTGT AAAAGAACAA ATCGATGAAA GTCT 954

## (2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:  
35 (A) LENGTH: 1602 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

50 CCTGATTATA CCCAACCTCT TTGCATCAAG TCGGAAAAAT GAGTGAAATG GGTTCAGT 60  
TTTCCTGAAA TAAGGTATCC TATATAAAGT ACCCTATGAT AACCATGGAG GTATTGTGTA 120  
55 TGGTTCAAAC AAGTCATTGA AGAAATACAA AACAATGCCA ACATTGTGGA AGTCATAGGA 180  
GATGTGATAT CTTACAAAAG GCAGGACGGA ACTATCTAGG GCTCTGTCCT TTTCATGGTG 240  
60 AAAAAACACC ATCTTTCAGC GTTGTAGAGA ACAAGCAGTT TTACCACTGT TTTGGTTGTG 300  
GTCGCTCAGG TGATGTCTTT AAAATTCATC GAGGAGTACC AAGGGGTAC CTTTATGGAG 360

-139-

5 GCTGTCCAAA TCTTAGGTCA GCGTGTCTGGG ATTGAGGTTG AAAAACCCTT TTATAGTGAA 420  
CAGAAGCCAG CCTCGCCTCA CCAAGCTCTT TATGATATGC ACGAAGATGC GGCTAAATTT 480  
TACCATGCTA TTCTCATGAC AACGACTATG GGCGAAGAGG CCAGAAATTA CCTTTATCAG 540  
CGGGGTTTGA CAGATGAAGT GCTTAAACAT TTTTGGATTG GTTTAGCACC TCCAGAACGA 600  
10 AACTATCTCT ATCAACGTTT GTCTGATCAG TATCGTGAAG AGGATTTACT GGATTCAGGC 660  
CTGTTTTATC TTTCGGATGC CAATCAATTT GTAGACACCT TTCACAATCG CATTATGTTT 720  
CCCCTGACAA ATGACCAAGG AAAGGTCATT GCCTTCTCAG GTCGTATCTG GCAAAAAACG 780  
15 GATTACAAA CTTCTAAGTA TAAAAACAGC CGTTCGACTG TAATTTTAA CAAAAGTTAC 840  
GAATTATATC ATATGGATAG GGCAAAAAGA TCTTCTGGAA AAGCTAGTGA GATTTACCTG 900  
20 ATGGAAGGAT TCATGGATGT TATTGCAGCC TATCGGGCTG GAATCGAAAA TGCTGTGGCG 960  
TCGATGGGAA CGGCCTTGAG TCGAGAGCAT GTTGAGCATC TGAAAAGGTT AACCAAGAAA 1020  
TTGGTTCTTG TTTACGATGG AGATAAGGCT GGGCAAGCCG CGACATTGAA AGCATTGGAT 1080  
25 GAAATTGGTG ATATGCCTGT GCAAATCGTC AGCATGCCTG ATAACCTGGA TCCTGATGAA 1140  
TATCTACAAA AAAATGGTCC AGAAGACTTG GCCTATCTAT TAACGAAAAC TCGTATTAGT 1200  
30 CCGATTGAGT TCTACATTCA TCAGTACAAA CCTGAAAACG GTGAAAATCT GCAGGCTCAG 1260  
ATTGAGTTTC TTGAAAAAAT AGCTCCCTTG ATTGTTCAAG AAAAGTCCAT CGCTGCTCAA 1320  
AACAGCTATA TTCATATTTT AGCTGACAGT CTGGCGTCCT TTGATTATAC CCAGATTGAG 1380  
35 CAGATTGTTA ATGAGAGTCG TCAGGTGCAA AGGCAGAATC GCATGGAAAG AATTTCCAGA 1440  
CCGACGCCAA TCACCATGCC TGTACCAAG CAGTTATCGG CTATTATGAG GGCAGAAGCC 1500  
40 CATCTACTCT ATCGGATGAT GGAATCCCT CTCGTTTTGA ACGATTACCG TTTGCGAGAA 1560  
GACTTTGCAT TTGCTACACC TGAATTCAG GTCTTACATG AC 1602

45 (2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7203 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

55 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:



-140-

	CCTCCATCAA ATCTGAGACT GATTCAAAAG ACTGGCTCAT ATTACGATTT TGGTCTAAAT	60
5	GCGTTAACAC TTGGAGCAAC TTCCGATTTT CGTCTAGTCT AACATCAAAA GGTAATCCCT	120
	GATATTGAAT TGCCTGACGA AAGGAAAATA TTAATAGCTG TTGTCATATC CATTCCCAGA	180
	TTACTAAACA CCTGTTGGGC CTGCTCCTTA ACCTCACTAT CCAGACGGAT GCTCATACTC	240
10	ATCTTTGACA TACTCTCACC CTCTTTCCAT AGACTATTTT AACAAAAAAG AAAGCTAATG	300
	TAAATCTATT GGATATACGT TAGCCTCTTC TAATAGATTA TTAAGCAATT TTTTAAAACA	360
15	ACTCATCAAA CAAACTCAAC TGTTATCCT CTGGCATATT TCCAAGAATA CCCATCTCAT	420
	CCATCTTTTC AACCAAGGTT GATGAGAGTC CACCACGCTT GCGTAGTTCT GTTTTAGAGA	480
	GGAATTCTCC CTCTTCACGC GCCCGCACCA GTTGCTTGGC AACGTTCTCT CCCAGACCAT	540
20	CCATTGCTAC AAATGGTGGG ATAAGGGTAT CCCCGTCGAT GAGGAACTCT GTCGCCTGAC	600
	TACGGTAGAG ATCTAATTG CCAAACCTGA AACCTCGTTC CCACATCTCA TTGACAATCT	660
25	CAAGAGTTGT ATAGAGATCG ATTTCCACAT TAGAGGCTTC ATTGTTCTTC CGTTTTTCAG	720
	AGATTTCTTC CATTCTGCGC TTGATGGCCT CCAAGCCCGC ACCCATGGTC TTGATATCAA	780
	AAGCCTTAGC ACGAATGGAG AAGTAAGCAC AGTAGTAATA AATAGGATGG TGAACCTTGA	840
30	AGTAAGCTAC ACGCAAGGCC ATCATAACGT AGGCTGCCGC ATGGGCCTTA GGAACATGT	900
	ACTTAATTTT CCCACAGGAT TCGATATACC ACTCTGGCAC CTTATTAGCC TTCATGGCTT	960
35	CGATATAGCC ATTTCTCTCC TCTTCTGAAA TCTTTAGCCA CAAACCCTTA CGTACCCGTT	1020
	CCATAATGGT AAAGGCCATC TTAGGTTCCA GACCCGCATG CATGAGGTAA ACCATGATGT	1080
	CGTCCCGACA ACCGATAACA GTCGATAGGT CCGCTATTCC TTGCTTAATC AGATCCTGAG	1140
40	CATTCCCCAA CCAAACGTCA GTACCGTGGG ACAGACCAGA CAGCTGAAGC AATTCCGCAA	1200
	AGGTTGTGG ATGGGTTTCG TCTACCATT CACGTACGAA ATTTGTTCCA AACTCTGGAA	1260
45	TCCCTAACAT ACCCGTAGCG TTCCAATTTG TTCAGGTGTT ACCCCTAGCA CATCAGTCCC	1320
	AGAAAAGAGT GCCATCACGC CTTGTCATC CATAGGAATT TTATTAGGGT CAATACCAGA	1380
	CAAATCCTGA AGTTTTCGAA TCATGGTCGG ATCATCATGT CCCAGTACAT CGAGTTTGAG	1440
50	GACGTTCTCA TCGATATCGT GGAAGTTAAA GTGAGTGGTC TGCCATTGAG CCGTGACATC	1500
	ATCTGCTGGA TACTGGACAG GCGTAAAATC GTAGACATCC ATGTAGTTCG GAATAACAAC	1560
55	GATTCCCCC GGGTGTGGC CTGTTGTCCG CTTGACACCC GCCGCTCCTT GAGCGAGGCG	1620
	TTCTACTTCT GCATCACGAT AAAACTTGCC ATAATCTCGC TCGTAACCCT TGACAAATCC	1680
	ATAGGCAGTC TTGGCAGCTA CCGTACCAAC TGTTCCCGCA CGGAAGGCAT ATTCTTCACC	1740
60	AAAGATATCA CGCACATCCA AGTGGGCGCT AGGCTGATCT TCTCCCGAGA AGTTCAAGTC	1800

-141-

	AATATCAGGA	ACCTTATCCC	CATCAAAACC	AAGGAAGGTC	TCAAACGGAA	TATCCTGTCC	1860
	GTCTTTACTG	AGTTTGTGAC	CACAGTTTGG	ACAGTCCTTA	TGGGGCATAT	CAAATCCTGA	1920
5	ACCGTACGAA	CCATCTGTGA	TAAACTCACT	GTACTGACAC	TGACCACAGA	CATAGTGAGG	1980
	AGAGAGAGGA	TTGACCTCCG	TAATCCCAAT	CATGGTCGCA	ACGAAACTAG	ATCCGACAGA	2040
10	CCCACGAGAA	CCAACCAAAT	AACCCCGTTC	ATTAGAACGT	TGCACCAGCA	TCTGCGATGC	2100
	CAGATAAATC	ACAGCAAATC	CATTCCCCAG	TATGGATGTT	AATTCTTTTT	CAATCCGCAA	2160
	ATCAACAATA	TCTGGCAGCG	GATTTCCATA	AATCTCAAAA	GCTTTCTTAT	AGGTCAAATC	2220
15	AGCAACTGTT	TCTTCAGCCT	TGTCGATGAA	AGGCGTATAC	AAGTCACCCCT	TAACGACTTC	2280
	AACGGGTTC	AATATTTCTG	CCAAGGCATT	GGTGTTC	ATAACCAGTT	TACGAGCCAG	2340
20	TTCTCTCCC	AAAAAGGCAA	ATTCATCCAA	CATCTCATT	GTCGTCGAA	AATGAGCCTT	2400
	TGGAAGTGGT	GCTGGTTGGG	CATGTTACC	ATGACCGATA	GTCGGTTAA	TCATCGCACC	2460
	CTGTCCCAA	CTACGGACGA	TAATTTACG	ATAAATCTCT	TCTCCGGTT	CGATATAGTG	2520
25	AACATTTCCC	GTAGCCAAA	CAGGCTTGCC	AAGGCGGTCT	CCAACCTCTA	TCAAATCTT	2580
	GATAATGGTC	TGGAGTTCCT	CCATATCCTT	GACCTGCTCT	TTAGCAATCA	AGGGCGCATA	2640
30	GATAGCCGGT	GGCATGACCT	CGATAAAGTC	ATAATACTTG	GCCACCTCAA	CCGCCGCATC	2700
	CACACCTTGA	GAAACGACCA	CGTCAAAAAC	TTCACCTCT	GAACAGGCTG	AACCTAAAT	2760
	CAAGCCCTCT	CGATGGGCAT	CTAGAACCGT	TCTCGGAATC	CGTGACACTC	CTTCAAAATA	2820
35	CTTGGTATTA	GACAAGGAAA	CCAGCTTAAA	GATATTTTTT	AGACCTACCT	GATTCTTGAC	2880
	ATAGATGGTC	GCATGCTTGA	TCCGAGCTTT	TTTGTAAGAA	TCTGGACTGA	TTAGATCAAT	2940
40	GTTGAGTCTA	GCTAAATCGG	TCACACCATG	TTTTTCTGCT	ACCTCTTTGA	TAAAGATAAA	3000
	GCAGACGACC	AGTCGCTTCC	GCATCGTATT	GGCCATGTGG	TGATGTTCCA	AGCCACACCA	3060
	AAACGCTTGG	TCAAAGGCCC	AAACCATGAT	TTATACTCAG	GATAGAGGTT	TCAGCAAATC	3120
45	CCAGGTATCA	ATAACGGCTG	ACTAATCTTT	GGCCATGACG	CTCATAATTA	GCATTCATAA	3180
	AGCCAACGTC	AAAGGTAGCA	TTGTGGGCAA	CTAGGACCGT	ATCCTTGCAA	ATTCTTGAA	3240
50	TTCTTGCAA	ACTTGTTCTA	GTGGTTTGGC	ATTTTGGACA	TGATCATCTG	TAATTCAGT	3300
	TAACTCTGTA	GTAAGGCTG	ACAAGGGATG	CCCAGGATTG	ATAAATTCAT	CAAATTCAGC	3360
	AATAACATTC	CCCTTGATCA	TCTTAGAGGC	CGCAACCTGA	ATCAAGTCAT	TATAGATAGC	3420
55	TGAAAGTCCC	GTCGTTTCCA	CGTCAAAGAC	CACGTAGGTT	GCTTCTGATA	AGTCCATCTA	3480
	CATTCTGTTAT	AGACGATAGG	GACACGGTCC	TCCACGATAT	TGGCTTCTAT	CCCATAGATC	3540
60	AGCTGGATT	CCGCTTTCTT	AGCCGCCTTA	TAGCCATGTG	GAAAGGACTG	GACATTCCCA	3600
	TGGTCTGTGA	TAGCAACCGC	CTTGTGTCCC	CACTTAGCAG	CTGTTGCAAC	AATCTCTTCG	3660

-142-

	ACCTCTGGCA AAGCATCCAT AGTCGACATG TTAGTATGAG CATGAACTC AACCCGACGC	3720
5	TCACCTTCTG GCATCAAATC CTTCCGCTCA TAGTGAACAA CTTCTGCAG ATCCTGTACG	3780
	TTCATAGTCA AATCGCGTGT GAAGTTATTC ATCTCCACAT TCCCTCGAAC TCGGAGCCAA	3840
	GAATTCTTCT TGATGAGGTC AAACCTCTGG GCCTCTTCCT CGTTTTTAAC CCACCTTTGC	3900
10	ATAGAAAAAC TTGAAGTATA GTCCGTCATT TTAAAGTTGA TTAAACACG ACCTGTCTA	3960
	GTCACCTTTT GCTCCACATC AAAACAACC CCTTCAAATA CCAGACGATT TTCCTCTGTC	4020
15	GTCACCTCGA TCATAGGAGT AATCTCCGCC TTATCCAGCT TGGGTTTAGC TGCAGCTTTT	4080
	TTGCGTTGAA AATCAAAGAC TGGTTTCTCT TCCGCTGGAG GAGGTGCCAT CTGCTCCAGT	4140
	TGTTCCATAG CACGGAGCGC TTCCTCATTG GCAGCTTGAA CAATCTGCTC ATTTTCAGCA	4200
20	TGAAAGGCCT CTTCTGCTC TTGGGTCAGG ACATCATTCT TCTCGACTTG ACAGTTAAAA	4260
	GTTGAAAAAC CAACTTTTC AAGTTGTTTG GCTAAATTAG GAAGATGATT CTTCTTAAAA	4320
25	TGTTCCCTAT CAATCGCTTC AGATCCTTCA ATAAATAGCT GATTACCCTC AGCACGAACT	4380
	TGCAAAATTT GATAAAGGGA CTTAAACCT TGAAGTACAC ATGGACCTTC AGAGAAAGCC	4440
	TCCCTATAGT AGGACTGCAA GAGCTGATTT GAAAATTCTT GAGACCGAGC CTTAATTTCA	4500
30	AAAACAGCTT TATGCTGT CTTAGAAAAT TCTTCGCTCA AACCTTCTT TAATTCTAAA	4560
	AAGATTCAA TCGGTAAAAT ATTAGAAAAT ACGAAATGAA ACTCCCATAC CTTACTAATT	4620
35	TTATGAACCA CAACTCGCTC AATATTGGCC TGTGCTAAAG CAGGAGCCTG TCTCATTTCA	4680
	GCAGGCATCC CCAATTGATT CATCAAAATT TCAAACTAT TTGACATTCA TTTTCTCAC	4740
	ATTATTCTTC TACTATTTTA CCATATTTAG AGGTATTTTC TAAAGACAAA AGGAAGCCAC	4800
40	TAAGTGACTT CTTCTAGAG TGAGGACGGA TTAGTCTTCA CCTTTATTTT TCTTAATAAT	4860
	TTCTTCTTGT ACTGACTTAG GTACATCTTC GTAGTGGTCA AATACCATCA TGAATGTACC	4920
45	ACGTCCTTGA GATGCAGAAC GAAGAACTGT TGCCTAACCG AACATTCAG CAAGTGGAAC	4980
	GTAAGCACGA ACGATTGGC TGTACCCTG TGCTTCCATA CCATCTACAC GTCCACGACG	5040
	AGCAGTTACG TGACCCATAA CATCACCAG GTTTTCTTCT GGAACAGTGA TTGTTACAAG	5100
50	CATCATTGGT TCAAGGATAG CTGGTTGTGC TGATTAGCA GCTTCTTTAA GGGAAAGTGA	5160
	AGCCGCAATC TTGAAGGCAG TTTCAGATGA GTCGACATCG TGATATGAAC CATCATAAAG	5220
55	CTTAGCTTTA ACGTCAACCA TTGGGTAACC TGCAAGAACA CCGTTAGCCA TAGATTCTAC	5280
	CAAACCTTTT TCAACCGCTG GGATAAATTC ACGAGGAACC ACACCACCGA CGATTGCGTT	5340
	TTCGAATTCG AATCCTTTAC CTTCTTCGTT TGGAGTAAAT TCAATCCATA CATCACCAG	5400
60	TTGACCTTTA CCACCAGACT GACGTTTGAA GAATCCGCGT GCTTGAGTAG AAGCGCGGAA	5460

-143-

	TGTTTCACGG TAAGATACTT GAGGAGCACC TACGTTTCGCT TCAACTTTGA ACTCACGACG	5520
	CATACGATCA ACAAGGACGT CAAGGTGAAG TTCACCCATA CCTGAGATAA CTGTTTCACC	5580
5	AGTTTCAACG TTTGTTTCAA CGCGGAATGT TGGATCTTCT TCAGCCAATT TTTGAAGGGC	5640
	GATACCCATC TTGTCTTGGT CAGCTTTAGA TTTTGGCTCA ACCATCAATT GGATAACTGG	5700
	TTCTGGAACG TTGATTGACT CAAGGATGAT TTTAGCTTTT TCATCTGTCA ATGAGTCACC	5760
10	AGTTGTAGTA TCTTTCAAAC CAACGGCAGC AGCGATATCA CCTGAGTAAA CAGTGTCTGAT	5820
	TTCTTGACGG CTGTTAGCGT GCATTTGAAG GATACGTCCG ATACGTTTAC GTTTACCTTT	5880
15	AGAAGTATTC AATACGTATG AACCTGATTG AAGAACACCT GAGTAAACAC GGAAGAATGT	5940
	CAAACGACCT ACGAATGGGT CAGTCATGAT CTTGAAGGCA AGAGCTGCAA ATGGCTCTTC	6000
	GTCAGATGCT GGACGAATTT CTTCAGCGTC TGTATCTGGG TTAATACCTT TGATTGCTGG	6060
20	GATGTCAAGT GGAAGTTGAA GGTAGTCGAT AACCGCATCA AGCATCAATT GAACACCTTT	6120
	GTTTTTGAAG GCTGAACCAC ACAATACTGG GAAGAATTCA ACGTTGATAG TCGCTTTACG	6180
25	GATACCAGCT TTCAATTCTT CGTTAGTGAT TTCTTCACCT TCGAGGTATT TCATCATCAA	6240
	TTCTTCGTCA GTTTCAGCAA CTGCTTCAAT CAATTTTTCA CGGTATTCTT GAGCTTGGTC	6300
	AAGGTATTCA GCTGGGATGT CTTCTTCAAG GATATCCGTA CCAAGGTCGT TAGTATAGAT	6360
30	TTCAGCTTTC ATCTTGATCA AGTCAATGAT ACCACGGAAG TCATCTTCAG AACCGATTGG	6420
	CAATTGGATT GGGTGTGCAT TTGCTTGAAG ACGATCGTGA AGTGTGCTTA CAGAGTAAAG	6480
35	GAAGTCAGCA CCGATTTTGT CCATTTTGTT GGCAAATACG ATACGTGGAA CTCCGTACTC	6540
	AGTTGCTTGA CGCCAACTG TTTCAAGTTG AGGCTCAACA CCTGATTGTG AGTCAAGAAC	6600
	GGTAACCGCA CCATCCAATA CACGAAGAGA ACGTTGTACT TCGATTGTGA AGTCCACGTG	6660
40	TCCTGGTGTG TCGATGATGT TTACGCGGTG GTTGTTCAT TGAGCTGTTG TCGCAGCAGA	6720
	TGTGATCGTG ATACCACGTT CTTGCTCTTG CTCCATCCAG TCCATTTGTG ACGCACCTTC	6780
45	GTGAGTTTCA CCGATTTTGT GGATTTTACC AGTGTAGTAA AGAATACGCT CAGTAGTTGT	6840
	TGTTTTACCG GCATCGACGT GAGCCATGAT ACCGATATTA CGAGTTTTTT CAAGTGAAAA	6900
	TTGCGGTGCC ATGAGGTTTG TTTCTCCTAT TTATTTTGA TTTCTATTCT ATTATAACAC	6960
50	GATTTTAATA AAAACGGATA GGCAGGACCT ACCCGTTCTC AATGTTTTCA TGCTATTGTT	7020
	GGTTTCAACT TACGAGATGG TAAGCTGAGT TATAGCTAAT ACTAATCGAT TTAGCTAATT	7080
55	TGAACCCGGG CTAAAGTTAG TTAGCCGATA TGAGCTGGAA CGGGATGCTG CGCGAAAAAG	7140
	ATAAACTCC TTGTATTCAT CGAATACTGC GTCAGTTTCC TATTTTCACC TTGCATCCTT	7200
60	ACC	7203

(2) INFORMATION FOR SEQ ID NO:80:

-144-

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 1581 base pairs  
    (B) TYPE: nucleic acid  
5      (C) STRANDEDNESS: single  
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10    (iii) HYPOTHETICAL: NO

        (iv) ANTI-SENSE: NO

15

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

20    GCATACAAAG GGCATCAAGA ATATCCGTGT TGCCACAAGC TGCGAGAAAG ATTTATGCCT      60  
      ACCGCCGTTA TGACCTTAAT GAATCTCCAA AGACCGCTTT AGACCTTATC ATCCCAGATT      120  
      TGTTTTGTGCA TATTTTGAAC CCTGCTGAAC GTGAAAGAAA ATTAAAGCGC GAAGGTGTAG      180  
25    AAGAATTATA TCTCCTTGAT TTAGTAGTC AATTTCGCTAG TCTCACTGCA CAAGAATTCT      240  
      TTGCAACTTA TATCAAGGCT ATGAATGCCA AAATTATTGT TGCAGGTTTT GATTATACAT      300  
      TTGGTCTGA CAAAAAACA GCAGAAGATT TAAAGGATTA CTTTGATGGA GAAGTTATCA      360  
30    TTGTTCACC TGTAGAAGAT GAGAAAGGAA AGATTAGTTC AACTCGTATC CGTCAAGCTA      420  
      TTTTAGATGG AAATGTGAAA GAAGCAGGAA AACTTTTGGG GGCACCGCTT CCATCAAGAG      480  
35    GTATGGTAGT TCATGGTAAT GCTCGTGGTC GTACAATTGG TTATCCGACA GCGAATTTAG      540  
      TGCTTTTAGA CCGTACTTAT ATGCCAGCAG ATGGCGTTTA TGTCGTTGAT GTTGAGATTC      600  
      AAAGACAGAA GTATCGTGCT ATGGCTAGTG TCGGGAAAAA TGTGACCTTT GATGGAGAAG      660  
40    AAGCACGTTT TGAAGTCAAT ATTTTGTATT TTAATCAAGA TATTTATGGG GAAACCGTCA      720  
      TGGTTTATTG GCTTGATCGC ATTCGTGATA TGACCAAATT TGA CTCAGTT GACCAATTAG      780  
45    TGGATCAGTT AAAGGCTGAT GAAGAAGTAA CTCGGAATTG GTCTTAAGAG CTTGAGTAAA      840  
      TAAAACAAAA AAGAGGTTGT CTGTAACCCA AAAGATAGAT GATTTAGTCT AACTTTTGAG      900  
      GTCACGACAT TACCTCTTTT TATTCTTTTT CAAAGGTGAA GCCTTCTCCT AGGATTTTATC      960  
50    GGGCTTCTGT AATAGTTATA AAGGCTTGAG GATCGATTCTG ATGAATCATT TCCTTCGTTT      1020  
      TCACAATTTT ATTTCTTCCG ACAATACAGT AGATGATTTT CAAATTTTCT TTA CTATAGT      1080  
55    AGCCTTGACC AGAAATAAAA GTAACACCTC TTCCGAGGTC ATCATTAATC GCCTTAGCAA      1140  
      GTTGGTCAGG ACGTTTTGTG ATAATCATAA AGCCTTTGCC GGCATATCCT CCTTACCAA      1200  
      TCAAATCAAT AACACGAGAA ACAATAAAAT CAAACAAAAG CGTGTAGGAA ACCAATCTCA      1260  
60    AATCCTTGAA GATTAGGAGA ATCAACATGA GAATACAAAA ATCTAAGATA AAGAGCAGTT      1320

-145-

5 TTCCTATGGA TATATGAGTG TATTTGTTGA GAATACGAGC TAGAATATCA GTTCCGCCAG 1380  
TTGTACCTCC AGCATTAAAA ATAATTCCAA GGCCAATTCC CAATAGGATT CCCGCTATAA 1440  
GGGCTGTGAT TAGTAAATCA CCTTGAAGAT CAATATGAAG GGGAATATGC TCAAAAAAAG 1500  
CTAACCAGGC GGACAAAGCT AAGGTTCCCTA GTAAACTAGA ATAGAGGGAT TTGGCTCCAA 1560  
10 AGATCTTCCA AGCTAGGATG A 1581

## (2) INFORMATION FOR SEQ ID NO:81:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 879 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

CAAGTTTGTC GAATTGCCAA ACACAGTTGA AGGCTTGATT CACATCACTA ATCTACCTGA 60  
ATTTTATCAT TTCAATGAGC GTGATTGAC TCTTCGTGGA GAAAAATCAG GTATCACTTT 120  
35 CCGAGTGGGT CAGCAGATCC GTATCCGTGT TGAAAGAGCG GATAAAATGA CTGGAGAGAT 180  
TGATTTTTC A TCGTACCTA GTGAGTTGA TGTGATTGAA AAAGGCTTGA AACAGTCTAG 240  
TCGTAGTGGC AGAGGGCGTG GTTCAAATCG TCGTTCGGAT AAGAAGGAAG ACAAGAGAAA 300  
40 ATCAGGACGC TCAAATGATA AGCGTAACAT TTCACAAAAA GACAAGAAGA AAAAAGGAAA 360  
GAAACCTTTT TACAAGGAAG TAGCTAAGAA AGGAGCCAAG CATGGCAAAG GGCGAGGGAA 420  
45 AGGTCGTGCG ACAAATAAAA AAGGCACGCC ACGACTATAC AATCGTAGAT ACGCTAGAGG 480  
CAGGGATGGT CCTGACTGGA ACTGAAATCA AGAGTGTACG AGCTGCTCGA ATTAATCTCA 540  
AGGATGGCTT TGCTCAAGTG AAAAATGGAG AAGTTTGGCT GAGTAATGTT CATATCGCGC 600  
50 CTTACGAAGA GGGCAATATC TGGAACCAGG AACCAGAACG TCGTCGTAAA CTCCTGCTCC 660  
ATAAAAAGCA AATTCAAAAA TTGGAACAAG AGACCAAAGG GACAGGAATG ACCTTAGTTC 720  
55 CCCTTAAGGT CTATATAGAT GGCTACGCTA AGCTTCTTTT AGGACTTGCC AAGGGAAGCA 780  
TGACTATGAC AAACGGAGTC TATCAAACGT CGTGAGCAAA TCGAGATATC GCGCGTGTGA 840  
TGAAGCTGTT AATCAGCGAT AAAGAGAGGA ATTGAGATG 879

## (2) INFORMATION FOR SEQ ID NO:82:

-146-

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1550 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

20	AAAAGCTTAA TAAATCAATA ATTTCTTCTT TTATCCCCAA CCTGTGGATA AAGTTTGGTA	60
	ACATTGTGGA TTATTTTTC AAGCTTGTGG AAAATCTTG CTATCTATGG TAAATATCT	120
	CTAGTATTAA ACTTTTAAAT AGTAAAGGAG GAGAAAGGAT TGAAAGAAAA ACAATTTTGG	180
25	AATCGTATAT TAGAATTGTC ACAAGAAAGA CTGACTCGAT CCATGTATGA TTTCTATGCT	240
	ATTCAAGCTG AACTTATCAA GGTAGAGGAA AATGTTGCCA CTATATTTCT ACCTCGCTCT	300
	GAAATGAAAA TGGTCTGGGA AAAACAATA AAAGATATTA TTGTAGTAGC TGGTTTGGAA	360
30	ATTTATGACG CTGAAATAAC TCCCCACTAT ATTTTCACCA AACCTCAAGA TACGACTAGC	420
	TCACAAGTTG AAGAAGCTAC AAATTAACT CTTTATGACT ATAGTCCAAA GTTAGTATCT	480
35	ATTCCTTATT CAGATACGGG ATTAAAAGAA AAGTATACCT TTGATAACTT TATTCAAGGG	540
	GATGGAAATG TTTGGGCTGT ATCAGCCGCT TTAGCTGTCT CTGAAGATTG GGCTCTGACC	600
	TATAACCTC TTTTATCTA TGGAGGACCA GGCCTTGGTA AGACTCACTT ATTAAACGCT	660
40	ATTGGAAATG AAATTCTAAA AAATATTCCT AATGCGCGTG TTAAATATAT CCCTGCCGAA	720
	AGCTTTATTA ATGACTTTCT TGATCACCTA AGACTTGGGG AAATGGAAAA GTTTAAAAAG	780
45	ACCTATCGTA GTCTTGATCT TTTGTTAATC GATGATATCC AGTCACTCAG CGGAAAAAAA	840
	GTCGCAACTC AGGAAGAATT TTTCAATACC TTTAACGCCC TTCATGACAA GCAAAAACAG	900
	ATTGTCCTAA CGAGTGATCG TAGTCCAAAA CATCTAGAAG GGCTCGAGGA GAGGCTTGTC	960
50	ACGCGTTTTA GTTGGGGATT GACACAACT ATCACACCCC CTGACTTTGA AACACGTATT	1020
	GCCATTTTAC AAAGTAAAC GGAACATTTA GGCTACAATT TCCAAAGTGA TACTCTAGAA	1080
55	TACCTAGCTG GGCAATTGTA TTCAAATGTT CGAGATCTTG AGGGAGCCAT CAACGACATC	1140
	ACTTTAATTG CCAGAGTAAA AAAAATCAAG GATATCACTA TTGATATTGC TGCAGAAGCC	1200
	ATTAGAGCCC GCAAACAAGA TGTTAGCCAA ATGCTCGTCA TCCCAATTGA TAAATCCAA	1260
60	ACTGAAGTTG GTAACTTTGA TGGTGTTAGT ATCAAAGAAA TGAAGGGAAG TAGACGCCTT	1320

-147-

CAAAATATTG TTTTGGCCCG TCAAGTAGCC ATGTATTTAT CTAGAGAACT AACAGATAAT 1380  
AGTCTTCCAA AAATTGGGAA GGAATTGGGG GAAAAGTCAT ACCACAGTCA TTCATGCCCA 1440  
5 TGCCAAAATA AAATCTTGAA TTGATCAAGA CGATAATTTA CGTTTAGAAA TTGAATCATC 1500  
AAAAGGAAAA TCAAATAATT TGTGGAACT TTAGGTTTTT ACCTTTTAGC 1550

10 (2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1292 base pairs  
(B) TYPE: nucleic acid  
15 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

30 GGTATGCGCC AAAACTTCTT ATCAAAAAGA ATCTAGCCAA AGAAGCGACT GCTCAAGCTG 60  
TAGGTGAACT TCGTGGTAAA CAAAATCGG AAGAAAAGC TCACGCTGAG ATGATTGCAG 120  
AAGGAAAAGC AATTAAAGCA CAACTTGAAG CAGAAGAAAC TGTGTAGAA TTTGTTGAAA 180  
35 AAGTTGGTCC AGATGGTCGT ACCTTTGGTT CTATTACCAA TAAGAAGATT GCAGAAGAAT 240  
TGCAAAAGCA ATTTGGAATT AAGATTGATA AACGTCATAT TCAAGTACAA GCTCCGATTC 300  
GAGCGGTTGG TTTGATTGAT GTGCCAGTGA AAATCTATCA AGATATCACA AGTGAATCA 360  
40 ATCTTCGTGT GAAAGAAGGA TAAGTTTACA CCTTCTTGAC AAGATTGTAA AAGGAAGGGA 420  
AGTCTGATGG CAGAAGTAGA AGAGTTACGA GTACAACCTC AAGATATCTT AGCTGAGCAA 480  
45 TCCGTTTTAG GGGCTATCTT TATTGATGAG AGTAACTTG TTTTGTGCG AGAATACATT 540  
GAGTCTCGGG ACTTTTTTAA GTATGCCCAT CGTTTGATT TCCAAGCCAT GGTCGATTTA 600  
TCCGATCGTG GTGATGCCAT AGATGCAACA ACGGTCGTA CTATCCTTGA TAATCAAGGT 660  
50 GATTACAGA ATATTGGTGG CTTGTCTTAC TTGGTTGAGA TTGTTAATTC TGTGCCAACT 720  
TCTGCTAATG CGGAGTATTA TGCTAAGATT GTTGCAAGAA AAGCAATGCT ACGTCGTTTA 780  
55 ATTGCCAAGT TGACAGAGTC TGTCAACCAA GCTTACGAAG CGTCACAACC AGCTGATGAA 840  
ATTATTGCTC AGGCAGAAAA AGGGTTGATT GATGTCAGTG AAAATGCAAA TCGAAGCGGG 900  
TTTAAGAACA TTCGAGATGT GTTGAATCTC AACTTTGGAA ATCTGGAAGC TCGCTCGCAA 960  
60 CAAACGACCG ATATTACAGG TATTGCGACA GGTATCGTG ATTTGGATCA TATGACAACA 1020



-148-

GGACTTCATG AGGAGGAGTT GATTATCTTA GCAGCTCGTC CAGCAGTTGG TAAGACAGCA 1080  
TTTGCCTTGA ATATCGCTCA GAATATTGGG ACTAAGTTGG AAAAAACGGT TGCTATTTTT 1140  
5 TCACTCGAAA TGGGTGCGGA AAGCTTGGTA GACCGTATGT TAGCTGCAGA AGGCTTGGTG 1200  
GAGTCACATT CTATCCGTAC AGGGCAATTG ACAGATGAGG AGTGGCAAAA ATATACTATT 1260  
10 GCTCAGGGTA ATCGTACTAA CGCCAGTATC TA 1292

## (2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:  
15 (A) LENGTH: 1876 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

30 AGGCTGATCC TCGGTTATT GACCCATGGA AGAGCACCTA GAAGGAGATC ATTCCCAGAC 60  
GATATTTTAG TTTTATCCTA GTAGCCTTCC CTGGCTATTT TAGGAGCTCG TCTCTACTAT 120  
35 GTATTTCCGA TTTGATTACT ATAGTCAGAA TTTAGGAGAG ATTTTGGCCA TTGGAATGGT 180  
GGTTGGCCAT TTACGGTGGT TGATAACTGG GGCTCTTG TGCTATATCT TTGCTGACCG 240  
TAAACTCATC AATACTGGG ATTTTCTAGA TATTGCGGCG CCTAGCGTTA TGATTGCTCA 300  
40 AAGTTTGGGG CGTTGGGGTA ATTTCTTTAA CCAAGAAGCT TATGGTGCAA CAGTGGATAA 360  
TCTGGATTAT CTACCTGGCT TTATCCGTGA CCAGATGTAT ATTGAGGGGA GCTACCGTCA 420  
45 ACCGACTTTC CTTTATGAGT CTCTATGGAA TCTGCTTGGC TTGCCTTGA TTCTGATTTT 480  
TAGACGGAAG TGGAAGAGTC TCAGACGAGG TCATATCACG GCCTTTTACT TGATTTGGTA 540  
TGTTTCGGT CGTATGGTCA TCGAAGGTAT GCGAACAGAT AGTCTCATGT TCTTCGGACT 600  
50 TCGAGTGTCC CAATGGCTGT CAGTTGTCCT TATCGGTCTC GGTATAATGA TCGTTATTTA 660  
TCAAAATCGA AAGAAGGCCC CTTACTATAT TACAGAGGAG GAAAATAAA TGTTAGAAGT 720  
55 TGATATATT CTTGTTGCCC TAGCTTTGAT TGTCTTTTGT GTCTATCTGA TCATTACTGT 780  
ACAAAAGCTT GGTCGTGTCA TCGATGAAAC AGAAAAGACG ATTAACCT TGACTTCAGA 840  
TGTGGATGTG ACCTTGCATC ACACCAATGA GTTGTGGCT AAGGTCAATG TCTTGGCAGA 900  
60 TGATATCAAT GTCAAGGTGG CTACGATTGA TCCACTCTTC AGTGCTGTTG CAGATTTATC 960

-149-

5 TCTATCTGTT TCAGACCTCA ATGACCATGC GCGTGTCTTG AGCAAGAAAG CTTCATCAGC 1020  
TGGTTCAAAA AACTCAAGA CTGGTGCAAG TCTGTCAGCT CTTCGTCTTG CAAGTAAATT 1080  
TTTCAAAAAA TAAAAAAGGA GAATCCTTAT GGGTAAATTA TCCTCAATCC TTTTAGGAAC 1140  
GGTTTCAGGT GCAGCTCTTG CCTTGTTTTT AACAAAGTGAT AAGGGCAAAC AAGTTTGCAG 1200  
10 TCAGGCTCAA GATTTTCTAG ATGATTGAG AGAAGATCCG GAGTATGCCA AGGAGCAAGT 1260  
CTGTGAAAAA CTGACAGAAG TTAAGGAGCA GGCTACAGAT TTTGTTCTGA AAACAAAAGA 1320  
ACAGGTTGAG TCAGGTGAAA TCACTGTGGA CAGTATACTT GCTCAAGCTA AATCCTATGC 1380  
15 TTTTCAAGCG ACAGAAGCAT CAAAAATCA ATTAAATAAT CTCAAGGAAC AATGGCAAGA 1440  
AAAAGCCGAA GCTCTTGATG ACTCAGAAGA GATTGTGATT CATATAACAG AAGAATAAAC 1500  
20 CATCACCATC TCCGGACGGA CTATGTATCT GGGGATGGTG ATTTTATCT GGAATCTAGT 1560  
CTTTGTGGTA TAATAATTAC TATGCAGAAA AAACCAACGT CAGCCTATGT GCACATCCCA 1620  
TTTTGTACCC AGATTGTGTA TTATTGTGAT TTTTCAAAGG TCTTCATCAA AAATCAGCCA 1680  
25 GTCGACAGCT ATTTAGAGCA TCTGCTGGAA GAGTTTCGTT CTTATGATAT TGAAAAGTTG 1740  
TCAACCCTTT ATATCGGTGG TGGAACACGA CAGCCCTGTC GGCTCCGCAA CTGGAGGTGT 1800  
30 TACTGAATGG CTTGACTAAA AACTTGGATT TGTCTGCTTG GAGAGTGACC ATTGAAGCCA 1860  
TCCAGGCGAT TTGGAA 1876

## (2) INFORMATION FOR SEQ ID NO:85:

35 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1574 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

45 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TTGGAAGATT TCCCACTTTC AGTGACCAAC CCATACGGTC GACTAAGCT CATGCTAGAG 60  
55 GAAATTTTGA CTGATATTTA CAAAGCAGAC TCAGAATGGA ATGTTGTCTT GCTTCGTTAC 120  
TTTAACCCAA TCGGAGTCCA TGAGAGTGGT GATTTGGGAG AAAATCCAAA CGGTATTCCA 180  
AACAATCTCT TGCCATATGT GACTCAAGTA GCCGTTGGAA AATTAGAGCA AGTGCAAGTG 240  
60 TTTGGAGACG ATTACGATAC GGAAGATGGA ACAGGTGTTT GTGACTATAT CCACGTTGTC 300

-150-

GATTTGGCTA AGGGTCACGT TGCAGCTTTG AAAAAAATCC AAAAAGGTTT AGGACTAAAC 360  
5 GTTTATAACC TTGGAAGTGG TAAAGGTTAC TCAGTTCTTG AAATTATCCA AAACATGGAA 420  
AAAGCGGTGG GATGTCCTAT TCCTTACCGC ATCGTAGAAC GTCGCCCAGG TGATATCGCT 480  
GCCTGCTACT CAGACCCAGC AAAGGCTAAA GCAGAACTCG GTTGGGAAGC AGAACTCGAC 540  
10 ATCACCCTAAA TGTGTGAAGG CCATGGCGTT GGCAGAGCAA GCATCCAAAT GGATTTGAAG 600  
ACTAAGATGA TGATTCAAT CATCGTCCCT TGTTTAACGA AGAGGAAGTA CTTCTCTTTT 660  
TTTATCAGGC TCTGGAAGCT TTAATTCCAG ATTTGGAAAC AAAATCGAGT ATGTCTTTGT 720  
15 CGATGATGGA TCAAGTGATG GGACCTTGGG ACTCTTAAAG GCCTATCGGG AGCAAAATCC 780  
GGCAGTCCAT TATATTTCTT TCTCTGAAA TTTTGGCAA GAAGCAGCCC TTTATGCAGG 840  
20 CTTGCAATAT GCGACAGGAG ATTTGGTGGT GGTGATGGAT GCAGACCTCC AAGATCCTCC 900  
TAGTATGTTG TTTGAGATGA AAAATGTACT AGACAAAAAT GTAGACTTGG ACTGCGTTGG 960  
GACACGGAGA ACTAGTCGGG AGGGAGAACC CTTCTTTCGC AGTTTCTGTG CTGTTCTCTT 1020  
25 TTATCGCCTC ATGCAAAAAA TCAGCCCACT AGCTCTGCCG TCGGGTGTCC GTGATTTTCG 1080  
TATGATGAGA AGGTCTGTGG TCGATGCCAT TTAAGCTTG ACTGAGTCCA ATCGTTTTTC 1140  
30 TAAGGGACTC TTTGCCTGGG TCGGCTTTAA AACCCTACTAT CTGGACTATC CAAATGTCGA 1200  
AAGGCAGGCT GGCAAGACCA GTTGGAGTTT TAGGCAACTC TTTTTTTACT CCATTGAAGG 1260  
GATTGTTAAT TTTTCAGATT TCCCTTTGAC TATAGCCTTT GTAGCTGGTC TCCTATCTTG 1320  
35 TTTTCTTTCT CTGCTGATGA CCTTTTTTGT TGTGGTTCGG ACCCTCATTT TGGGCAATCC 1380  
GACATCTGGT TGGACCTCTC TGATGGCTGT TATTCTCTAT CTTGGAGGCA TTCAACTCTT 1440  
40 GACCATTGGG ATTCTCGGTA AGTATAATCA GTAAGATTAT TTAGAGACTA AAAAAAGAC 1500  
CACTTTATCT TATCAAAGAA AAAGTGACCT TCCTGATTTT ACAGAAACCT AAAGTGAAAA 1560  
GACTATAATT TTCC 1574

(2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:  
50 (A) LENGTH: 969 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)  
55 (iii) HYPOTHETICAL: NO  
(iv) ANTI-SENSE: NO  
60

-151-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

5	GTTATAATTA TTGATGATAA TTATAGTAAT GTAAATTAA GAAATAAAAT TATCCATCAA	60
	TTTGGCTATA CCAATCATAG AATTAAGTTA ATTTTAAGTA ATGAAGATTT AGGTGCAACT	120
	AATGCCAGAA ACATAGGTAT CAAAAATTCT AGAGGTAAGT ATATATCATT TTTAGACGAT	180
10	GATGATGAAT ATATGCCAGA TCGAATTTTA AAGTTGATGG CTTGTTTAA AAAGAGTAGA	240
	ATGAAGAATT TAGCTTTAGT TTATAGTTAT GGCATAATAA TTTATCCAAA TGGTACACGA	300
	GAAGAGGAGA AGACCGATTT TGTGGGAAAT CCCTTGTTTG TTCAAATGGT TCACAATATA	360
15	GCAGGTACGT CATTTTGGTT GTGTAAAAA GAGGTGCTAG AATTAATTA TGGTTTGGAG	420
	AAAATAGATT CACATCAGGA CGGTGTTGTT TTATTAAAC TACTTGCTCA AGGATACCAA	480
20	ATTGATATAG TGCAGAGAATT CTTGGTGAAT TACTACGCTC ACAGTAAAGA AAACGGTATC	540
	ACTGGAGTGA CACAAAAAC AATTAATGCA GATGAAGAAT ATTATAATTA CTGTAGGAAA	600
	TATTTTAATT TATTGAGTTT CAACGAGAGA ATATTGGTTA CAAAGAAATA TTATTCTTTA	660
25	AACATAAAGC GGTACTATT AATAGGAGAC AAATGCAAGG CTTTAAAGT AATCAAGAAG	720
	GCAAGAGAAG AAAAAATTT TAACGAATTT CTTTTTTGA AATATATGTT ATTATATAAC	780
30	GTAGTTTTTT CTATTGTATA TATGACAACT ATGTTCAATT AAAATTTAGA AAGTGAGAAA	840
	CTATTGTGTA TACTATTATA AATTCAATAT AAACATTTAG GTTAATTAAC GATAATTAAT	900
	CGGTGCTGGG TCATTAATTG CTAATTTAAT GCAGCACTAT TAATGCTCAG GTGTTGAATG	960
35	AATTAATGC	969

## (2) INFORMATION FOR SEQ ID NO:87:

40	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 1353 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
	(D) TOPOLOGY: linear
45	(ii) MOLECULE TYPE: DNA (genomic)
	(iii) HYPOTHETICAL: NO
50	(iv) ANTI-SENSE: NO
	(ix) FEATURE:
	(A) NAME/KEY: CDS
55	(B) LOCATION: 1..1350
	(D) OTHER INFORMATION: DNA B

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

60	ATG GCA GAA GTA GAA GAG TTA CGA GTA CAA CCT CAA GAT ATC TTA GCT	48
	Met Ala Glu Val Glu Glu Leu Arg Val Gln Pro Gln Asp Ile Leu Ala	

-152-

	1	5	10	15	
5	GAG CAA TCC GTT TTA GGG GCT ATC TTT ATT GAT GAG AGT AAA CTT GTT Glu Gln Ser Val Leu Gly Ala Ile Phe Ile Asp Glu Ser Lys Leu Val	20	25	30	96
10	TTT GTG CGA GAA TAC ATT GAG TCT CGG GAC TTT TTT AAG TAT GCC CAT Phe Val Arg Glu Tyr Ile Glu Ser Arg Asp Phe Phe Lys Tyr Ala His	35	40	45	144
15	CGT TTG ATT TTC CAA GCC ATG GTC GAT TTA TCC GAT CGT GGT GAT GCC Arg Leu Ile Phe Gln Ala Met Val Asp Leu Ser Asp Arg Gly Asp Ala	50	55	60	192
20	ATA GAT GCA ACA ACG GTT CGT ACT ATC CTT GAT AAT CAA GGT GAT TTA Ile Asp Ala Thr Thr Val Arg Thr Ile Leu Asp Asn Gln Gly Asp Leu	65	70	75	240
25	CAG AAT ATT GGT GGC TTG TCT TAC TTG GTT GAG ATT GTT AAT TCT GTG Gln Asn Ile Gly Glu Ser Tyr Leu Val Glu Ile Val Asn Ser Val	85	90	95	288
30	CCA ACT TCT GCT AAT GCG GAG TAT TAT GCT AAG ATT GTT GCA GAA AAA Pro Thr Ser Ala Asn Ala Glu Tyr Tyr Ala Lys Ile Val Ala Glu Lys	100	105	110	336
35	GCA ATG CTA CGT CGT TTA ATT GCC AAG TTG ACA GAG TCT GTC AAC CAA Ala Met Leu Arg Arg Leu Ile Ala Lys Leu Thr Glu Ser Val Asn Gln	115	120	125	384
40	GCT TAC GAA GCG TCA CAA CCA GCT GAT GAA ATT ATT GCT CAG GCA GAA Ala Tyr Glu Ala Ser Gln Pro Ala Asp Glu Ile Ile Ala Gln Ala Glu	130	135	140	432
45	AAA GGG TTG ATT GAT GTC AGT GAA AAT GCA AAT CGA AGC GGG TTT AAG Lys Gly Leu Ile Asp Val Ser Glu Asn Ala Asn Arg Ser Gly Phe Lys	145	150	155	480
50	AAC ATT CGA GAT GTG TTG AAT CTC AAC TTT GGA AAT CTG GAA GCT CGC Asn Ile Arg Asp Val Leu Asn Leu Asn Phe Gly Asn Leu Glu Ala Arg	165	170	175	528
55	TCG CAA CAA ACG ACC GAT ATT ACA GGT ATT GCG ACA GGT TAT CGT GAT Ser Gln Gln Thr Thr Asp Ile Thr Gly Ile Ala Thr Gly Tyr Arg Asp	180	185	190	576
60	TTG GAT CAT ATG ACA ACA GGA CTT CAT GAG GAG GAG TTG ATT ATC TTA Leu Asp His Met Thr Thr Gly Leu His Glu Glu Glu Leu Ile Ile Leu	195	200	205	624
65	GCA GCT CGT CCA GCA GTT GGT AAG ACA GCA TTT GCC TTG AAT ATC GCT Ala Ala Arg Pro Ala Val Gly Lys Thr Ala Phe Ala Leu Asn Ile Ala	210	215	220	672
70	CAG AAT ATT GGG ACT AAG TTG GAC AAA ACG GTT GCT ATT TTT TCA CTC Gln Asn Ile Gly Thr Lys Leu Asp Lys Thr Val Ala Ile Phe Ser Leu	225	230	235	720
75	GAA ATG GGT GCG GAA AGC TTG GTA GAC CGT ATG TTA GCT GCA GAA GGC Glu Met Gly Ala Glu Ser Leu Val Asp Arg Met Leu Ala Ala Glu Gly	245	250	255	768

-153-

	TTG GTG GAG TCA CAT TCT ATC CGT ACA GGG CAA TTG ACA GAT GAG GAG	816
	Leu Val Glu Ser His Ser Ile Arg Thr Gly Gln Leu Thr Asp Glu Glu	
	260 265 270	
5	TGG CAA AAA TAT ACT ATT GCT CAG GGT AAT CTA GCT AAC GCC AGT ATC	864
	Trp Gln Lys Tyr Thr Ile Ala Gln Gly Asn Leu Ala Asn Ala Ser Ile	
	275 280 285	
10	TAT ATC GAT GAT ACG CCA GGT ATT CGG ATT ACA GAG ATT CGT TCT CGT	912
	Tyr Ile Asp Asp Thr Pro Gly Ile Arg Ile Thr Glu Ile Arg Ser Arg	
	290 295 300	
15	TCT CGT AAA TTG GCT CAA GAA ACT GGA AAT CTT GGT TTG ATT TTG ATA	960
	Ser Arg Lys Leu Ala Gln Glu Thr Gly Asn Leu Gly Leu Ile Leu Ile	
	305 310 315 320	
20	GAC TAT TTG CAA CTT ATC ACG GGA ACT GGT CGA GAA AAT CGT CAA CAA	1008
	Asp Tyr Leu Gln Leu Ile Thr Gly Thr Gly Arg Glu Asn Arg Gln Gln	
	325 330 335	
	GAA GTT TCT GAA ATT TCT CGT CAG TTG AAA ATA CTA GCC AAG GAA TTG	1056
	Glu Val Ser Glu Ile Ser Arg Gln Leu Lys Ile Leu Ala Lys Glu Leu	
	340 345 350	
25	AAG GTT CCA GTA ATC GCT CTG AGT CAG CTT TCT CGT GGT GTA GAA CAA	1104
	Lys Val Pro Val Ile Ala Leu Ser Gln Leu Ser Arg Gly Val Glu Gln	
	355 360 365	
30	CGT CAG GAC AAG AGA CCG GTC TTG TCT GAT ATT CGT GAA TCT GGG TCT	1152
	Arg Gln Asp Lys Arg Pro Val Leu Ser Asp Ile Arg Glu Ser Gly Ser	
	370 375 380	
35	ATT GAG CAG GAC GCT GAT ATC GTA GCT TTT CTC TAT CGC GAT GAC TAC	1200
	Ile Glu Gln Asp Ala Asp Ile Val Ala Phe Leu Tyr Arg Asp Asp Tyr	
	385 390 395 400	
40	TAT GAA CGT GGT GGT GAA GAA GAG GAG GGT ATC CCA AAT AAT AAG GTG	1248
	Tyr Glu Arg Gly Glu Glu Glu Glu Gly Ile Pro Asn Asn Lys Val	
	405 410 415	
	GAA GTT ATT ATC GAG AAA AAC CGT AGT GGA GCT CGT GGA ACA GTG GAA	1296
	Glu Val Ile Ile Glu Lys Asn Arg Ser Gly Ala Arg Gly Thr Val Glu	
	420 425 430	
45	TTG ATT TTC CAA AAA GAA TAC AAT AAA TTT TCA AGT ATC TCA AAG AGG	1344
	Leu Ile Phe Gln Lys Glu Tyr Asn Lys Phe Ser Ser Ile Ser Lys Arg	
	435 440 445	
50	GAG GCA TAA	1353
	Glu Ala	
	450	

55 (2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 450 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

-154-

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

5 Met Ala Glu Val Glu Glu Leu Arg Val Gln Pro Gln Asp Ile Leu Ala  
    1                  5                  10                  15  
   Glu Gln Ser Val Leu Gly Ala Ile Phe Ile Asp Glu Ser Lys Leu Val  
                   20                  25                  30  
 10 Phe Val Arg Glu Tyr Ile Glu Ser Arg Asp Phe Phe Lys Tyr Ala His  
                   35                  40                  45  
   Arg Leu Ile Phe Gln Ala Met Val Asp Leu Ser Asp Arg Gly Asp Ala  
 15                  50                  55                  60  
   Ile Asp Ala Thr Thr Val Arg Thr Ile Leu Asp Asn Gln Gly Asp Leu  
                   65                  70                  75                  80  
 20 Gln Asn Ile Gly Gly Leu Ser Tyr Leu Val Glu Ile Val Asn Ser Val  
                   85                  90                  95  
   Pro Thr Ser Ala Asn Ala Glu Tyr Tyr Ala Lys Ile Val Ala Glu Lys  
                   100                  105                  110  
 25 Ala Met Leu Arg Arg Leu Ile Ala Lys Leu Thr Glu Ser Val Asn Gln  
                   115                  120                  125  
   Ala Tyr Glu Ala Ser Gln Pro Ala Asp Glu Ile Ile Ala Gln Ala Glu  
 30                  130                  135                  140  
   Lys Gly Leu Ile Asp Val Ser Glu Asn Ala Asn Arg Ser Gly Phe Lys  
                   145                  150                  155                  160  
 35 Asn Ile Arg Asp Val Leu Asn Leu Asn Phe Gly Asn Leu Glu Ala Arg  
                   165                  170                  175  
   Ser Gln Gln Thr Thr Asp Ile Thr Gly Ile Ala Thr Gly Tyr Arg Asp  
                   180                  185                  190  
 40 Leu Asp His Met Thr Thr Gly Leu His Glu Glu Glu Leu Ile Ile Leu  
                   195                  200                  205  
   Ala Ala Arg Pro Ala Val Gly Lys Thr Ala Phe Ala Leu Asn Ile Ala  
 45                  210                  215                  220  
   Gln Asn Ile Gly Thr Lys Leu Asp Lys Thr Val Ala Ile Phe Ser Leu  
                   225                  230                  235                  240  
 50 Glu Met Gly Ala Glu Ser Leu Val Asp Arg Met Leu Ala Ala Glu Gly  
                   245                  250                  255  
   Leu Val Glu Ser His Ser Ile Arg Thr Gly Gln Leu Thr Asp Glu Glu  
                   260                  265                  270  
 55 Trp Gln Lys Tyr Thr Ile Ala Gln Gly Asn Leu Ala Asn Ala Ser Ile  
                   275                  280                  285  
   Tyr Ile Asp Asp Thr Pro Gly Ile Arg Ile Thr Glu Ile Arg Ser Arg  
 60                  290                  295                  300

-155-

Ser Arg Lys Leu Ala Gln Glu Thr Gly Asn Leu Gly Leu Ile Leu Ile  
305 310 315 320

5 Asp Tyr Leu Gln Leu Ile Thr Gly Thr Gly Arg Glu Asn Arg Gln Gln  
325 330 335

Glu Val Ser Glu Ile Ser Arg Gln Leu Lys Ile Leu Ala Lys Glu Leu  
340 345 350

10 Lys Val Pro Val Ile Ala Leu Ser Gln Leu Ser Arg Gly Val Glu Gln  
355 360 365

Arg Gln Asp Lys Arg Pro Val Leu Ser Asp Ile Arg Glu Ser Gly Ser  
370 375 380

15 Ile Glu Gln Asp Ala Asp Ile Val Ala Phe Leu Tyr Arg Asp Asp Tyr  
385 390 395 400

Tyr Glu Arg Gly Gly Glu Glu Glu Glu Gly Ile Pro Asn Asn Lys Val  
405 410 415

Glu Val Ile Ile Glu Lys Asn Arg Ser Gly Ala Arg Gly Thr Val Glu  
420 425 430

25 Leu Ile Phe Gln Lys Glu Tyr Asn Lys Phe Ser Ser Ile Ser Lys Arg  
435 440 445

Glu Ala  
450

30

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1785 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: DNA (genomic)

40

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..1782  
 (D) OTHER INFORMATION: DNA G

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

ATG ATA ACC ATG GAG GTA TTG TGT ATG GTT GAC AAA CAA GTC ATT GAA 48  
 55 Met Ile Thr Met Glu Val Leu Cys Met Val Asp Lys Gln Val Ile Glu  
 1 5 10 15

GAA ATC AAA AAC AAT GCC AAC ATT GTG GAA GTC ATA GGA GAT GTG ATT 96  
 60 Glu Ile Lys Asn Asn Ala Asn Ile Val Glu Val Ile Gly Asp Val Ile  
 20 25 30



-156-

	TCT TTA CAA AAG GCA GGA CGG AAC TAT CTA GGG CTC TGT CCT TTT CAT	144
	Ser Leu Gln Lys Ala Gly Arg Asn Tyr Leu Gly Leu Cys Pro Phe His	
	35 40 45	
5	GGT GAA AAA ACA CCT TCT TTC AGC GTT GTA GAG GAC AAG CAG TTT TAC	192
	Gly Glu Lys Thr Pro Ser Phe Ser Val Val Glu Asp Lys Gln Phe Tyr	
	50 55 60	
10	CAC TGT TTT GGT TGT GGT CGC TCA GGT GAT GTC TTT AAA TTC ATC GAG	240
	His Cys Phe Gly Cys Gly Arg Ser Gly Asp Val Phe Lys Phe Ile Glu	
	65 70 75 80	
15	GAG TAC CAA GGG GTT ACC TTT ATG GAG GCT GTC CAA ATC TTA GGT CAG	288
	Glu Tyr Gln Gly Val Thr Phe Met Glu Ala Val Gln Ile Leu Gly Gln	
	85 90 95	
20	CGT GTC GGG ATT GAG GTT GAA AAA CCG CTT TAT AGT GAA CAG AAG CCA	336
	Arg Val Gly Ile Glu Val Glu Lys Pro Leu Tyr Ser Glu Gln Lys Pro	
	100 105 110	
25	GCC TCG CCT CAC CAA GCT CTT TAT GAT ATG CAC GAA GAT GCG GCT AAA	384
	Ala Ser Pro His Gln Ala Leu Tyr Asp Met His Glu Asp Ala Ala Lys	
	115 120 125	
30	TTT TAC CAT GCT ATT CTC ATG ACA ACG ACT ATG GGC GAA GAG GCC AGA	432
	Phe Tyr His Ala Ile Leu Met Thr Thr Met Gly Glu Glu Ala Arg	
	130 135 140	
35	AAT TAC CTT TAT CAG CGG GGT TTG ACA GAT GAA GTG CTT AAA CAT TTT	480
	Asn Tyr Leu Tyr Gln Arg Gly Leu Thr Asp Glu Val Leu Lys His Phe	
	145 150 155 160	
40	TGG ATT GGT TTA GCA CCT CCA GAA CGA AAC TAT CTC TAT CAA CGT TTG	528
	Trp Ile Gly Leu Ala Pro Pro Glu Arg Asn Tyr Leu Tyr Gln Arg Leu	
	165 170 175	
45	TCT GAT CAG TAT CGT GAA GAG GAT TTA CTG GAT TCA GGC CTG TTT TAT	576
	Ser Asp Gln Tyr Arg Glu Glu Asp Leu Leu Asp Ser Gly Leu Phe Tyr	
	180 185 190	
50	CTT TCG GAT GCC AAT CAA TTT GTA GAC ACC TTT CAC AAT CGC ATT ATG	624
	Leu Ser Asp Ala Asn Gln Phe Val Asp Thr Phe His Asn Arg Ile Met	
	195 200 205	
55	TTT CCC CTG ACA AAT GAC CAA GGA AAG GTC ATT GCC TTC TCA GGT CGT	672
	Phe Pro Leu Thr Asn Asp Gln Gly Lys Val Ile Ala Phe Ser Gly Arg	
	210 215 220	
60	ATC TGG CAA AAA ACG GAT TCA CAA ACT TCT AAG TAT AAA AAC AGC CGA	720
	Ile Trp Gln Lys Thr Asp Ser Gln Thr Ser Lys Tyr Lys Asn Ser Arg	
	225 230 235 240	
65	TCG ACT GTA ATT TTT AAC AAA AGT TAC GAA TTA TAT CAT ATG GAT AGG	768
	Ser Thr Val Ile Phe Asn Lys Ser Tyr Glu Leu Tyr His Met Asp Arg	
	245 250 255	
70	GCA AAA AGA TCT TCT GGA AAA GCT AGT GAG ATT TAC CTG ATG GAA GGA	816
	Ala Lys Arg Ser Ser Gly Lys Ala Ser Glu Ile Tyr Leu Met Glu Gly	
	260 265 270	
75	TTC ATG GAT GTT ATT GCA GCC TAT CGG GCT GGA ATC GAA AAT GCT GTG	864

-157-

	Phe	Met	Asp	Val	Ile	Ala	Ala	Tyr	Arg	Ala	Gly	Ile	Glu	Asn	Ala	Val	
			275					280					285				
5	GCG	TCG	ATG	GGA	ACG	GCC	TTG	AGT	CGA	GAG	CAT	GTT	GAG	CAT	CTG	AAA	912
	Ala	Ser	Met	Gly	Thr	Ala	Leu	Ser	Arg	Glu	His	Val	Glu	His	Leu	Lys	
			290				295					300					
10	AGG	TTA	ACC	AAG	AAA	TTG	GTT	CTT	GTT	TAC	GAT	GGA	GAT	AAG	GCT	GGG	960
	Arg	Leu	Thr	Lys	Lys	Leu	Val	Leu	Val	Tyr	Asp	Gly	Asp	Lys	Ala	Gly	
	305					310					315					320	
15	CAA	GCC	GCG	ACA	TTG	AAA	GCA	TTG	GAT	GAA	ATT	GGT	GAT	ATG	CCT	GTG	1008
	Gln	Ala	Ala	Thr	Leu	Lys	Ala	Leu	Asp	Glu	Ile	Gly	Asp	Met	Pro	Val	
					325					330					335		
	CAA	ATC	GTC	AGC	ATG	CCT	GAT	AAC	TTG	GAT	CCT	GAT	GAA	TAT	CTA	CAA	1056
	Gln	Ile	Val	Ser	Met	Pro	Asp	Asn	Leu	Asp	Pro	Asp	Glu	Tyr	Leu	Gln	
				340					345					350			
20	AAA	AAT	GGT	CCA	GAA	GAC	TTG	GCC	TAT	CTA	TTA	ACG	AAA	ACT	CGT	ATT	1104
	Lys	Asn	Gly	Pro	Glu	Asp	Leu	Ala	Tyr	Leu	Leu	Thr	Lys	Thr	Arg	Ile	
			355					360					365				
25	AGT	CCG	ATT	GAG	TTC	TAC	ATT	CAT	CAG	TAC	AAA	CCT	GAA	AAC	GGT	GAA	1152
	Ser	Pro	Ile	Glu	Phe	Tyr	Ile	His	Gln	Tyr	Lys	Pro	Glu	Asn	Gly	Glu	
		370					375					380					
30	AAT	CTG	CAG	GCT	CAG	ATT	GAG	TTT	CTT	GAA	AAA	ATA	GCT	CCC	TTG	ATT	1200
	Asn	Leu	Gln	Ala	Gln	Ile	Glu	Phe	Leu	Glu	Lys	Ile	Ala	Pro	Leu	Ile	
	385					390					395					400	
35	GTT	CAA	GAA	AAG	TCC	ATC	GCT	GCT	CAA	AAC	AGC	TAT	ATT	CAT	ATT	TTA	1248
	Val	Gln	Glu	Lys	Ser	Ile	Ala	Ala	Gln	Asn	Ser	Tyr	Ile	His	Ile	Leu	
					405					410					415		
	GCT	GAC	AGT	CTG	GCG	TCC	TTT	GAT	TAT	ACC	CAG	ATT	GAG	CAG	ATT	GTT	1296
	Ala	Asp	Ser	Leu	Ala	Ser	Phe	Asp	Tyr	Thr	Gln	Ile	Glu	Gln	Ile	Val	
				420					425					430			
40	AAT	GAG	AGT	CGT	CAG	GTG	CAA	AGG	CAG	AAT	CGC	ATG	GAA	AGA	ATT	TCC	1344
	Asn	Glu	Ser	Arg	Gln	Val	Gln	Arg	Gln	Asn	Arg	Met	Glu	Arg	Ile	Ser	
			435					440					445				
45	AGA	CCG	ACG	CCA	ATC	ACC	ATG	CCT	GTC	ACC	AAG	CAG	TTA	TCG	GCT	ATT	1392
	Arg	Pro	Thr	Pro	Ile	Thr	Met	Pro	Val	Thr	Lys	Gln	Leu	Ser	Ala	Ile	
		450					455					460					
50	ATG	AGG	GCA	GAA	GCC	CAT	CTA	CTC	TAT	CGG	ATG	ATG	GAA	TCC	CCT	CTT	1440
	Met	Arg	Ala	Glu	Ala	His	Leu	Leu	Tyr	Arg	Met	Met	Glu	Ser	Pro	Leu	
	465					470					475					480	
55	GTT	TTG	AAC	GAT	TAC	CGT	TTG	CGA	GAA	GAC	TTT	GCA	TTT	GCT	ACA	CCT	1488
	Val	Leu	Asn	Asp	Tyr	Arg	Leu	Arg	Glu	Asp	Phe	Ala	Phe	Ala	Thr	Pro	
					485					490					495		
	GAA	TTT	CAG	GTC	TTA	CAT	GAC	TTG	CTT	GGC	CAG	TAT	GGA	AAT	CTT	CCT	1536
	Glu	Phe	Gln	Val	Leu	His	Asp	Leu	Leu	Gly	Gln	Tyr	Gly	Asn	Leu	Pro	
				500					505					510			
60	CCA	GAA	GTT	TTA	GCA	GAG	CAG	ACA	GAG	GAA	GTT	GAA	AGA	GCT	TGG	TAC	1584
	Pro	Glu	Val	Leu	Ala	Glu	Gln	Thr	Glu	Glu	Val	Glu	Arg	Ala	Trp	Tyr	

-158-

	515	520	525	
5	CAA GTT TTA GCT CAG GAT TTG CCT GCT GAG ATA TCG CCG CAG GAA CTT Gln Val Leu Ala Gln Asp Leu Pro Ala Glu Ile Ser Pro Gln Glu Leu 530 535 540			1632
10	AGT GAA GTA GAG ATG ACT CGA AAC AAG GCT CTC TTG AAT CAG GAC AAT Ser Glu Val Glu Met Thr Arg Asn Lys Ala Leu Leu Asn Gln Asp Asn 545 550 555 560			1680
	ATG AGA ATC AAA AAG AAG GTG CAG GAA GCT AGC CAT GTA GGA GAT ACA Met Arg Ile Lys Lys Lys Val Gln Glu Ala Ser His Val Gly Asp Thr 565 570 575			1728
15	GAT ACA GCC CTA GAA GAA TTG GAA CGT TTA ATT TCC CAA AAG AGA AGA Asp Thr Ala Leu Glu Glu Leu Glu Arg Leu Ile Ser Gln Lys Arg Arg 580 585 590			1776
20	ATG GAG TAA Met Glu			1785

## (2) INFORMATION FOR SEQ ID NO:90:

25

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 594 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

35	Met Ile Thr Met Glu Val Leu Cys Met Val Asp Lys Gln Val Ile Glu 1 5 10 15
40	Glu Ile Lys Asn Asn Ala Asn Ile Val Glu Val Ile Gly Asp Val Ile 20 25 30
	Ser Leu Gln Lys Ala Gly Arg Asn Tyr Leu Gly Leu Cys Pro Phe His 35 40 45
45	Gly Glu Lys Thr Pro Ser Phe Ser Val Val Glu Asp Lys Gln Phe Tyr 50 55 60
	His Cys Phe Gly Cys Gly Arg Ser Gly Asp Val Phe Lys Phe Ile Glu 65 70 75 80
50	Glu Tyr Gln Gly Val Thr Phe Met Glu Ala Val Gln Ile Leu Gly Gln 85 90 95
	Arg Val Gly Ile Glu Val Glu Lys Pro Leu Tyr Ser Glu Gln Lys Pro 100 105 110
55	Ala Ser Pro His Gln Ala Leu Tyr Asp Met His Glu Asp Ala Ala Lys 115 120 125
60	Phe Tyr His Ala Ile Leu Met Thr Thr Thr Met Gly Glu Glu Ala Arg 130 135 140

-159-

Asn Tyr Leu Tyr Gln Arg Gly Leu Thr Asp Glu Val Leu Lys His Phe  
 145 150 155 160  
 5 Trp Ile Gly Leu Ala Pro Pro Glu Arg Asn Tyr Leu Tyr Gln Arg Leu  
 165 170 175  
 Ser Asp Gln Tyr Arg Glu Glu Asp Leu Leu Asp Ser Gly Leu Phe Tyr  
 180 185 190  
 10 Leu Ser Asp Ala Asn Gln Phe Val Asp Thr Phe His Asn Arg Ile Met  
 195 200 205  
 Phe Pro Leu Thr Asn Asp Gln Gly Lys Val Ile Ala Phe Ser Gly Arg  
 210 215 220  
 15 Ile Trp Gln Lys Thr Asp Ser Gln Thr Ser Lys Tyr Lys Asn Ser Arg  
 225 230 235 240  
 Ser Thr Val Ile Phe Asn Lys Ser Tyr Glu Leu Tyr His Met Asp Arg  
 245 250 255  
 Ala Lys Arg Ser Ser Gly Lys Ala Ser Glu Ile Tyr Leu Met Glu Gly  
 260 265 270  
 25 Phe Met Asp Val Ile Ala Ala Tyr Arg Ala Gly Ile Glu Asn Ala Val  
 275 280 285  
 Ala Ser Met Gly Thr Ala Leu Ser Arg Glu His Val Glu His Leu Lys  
 290 295 300  
 30 Arg Leu Thr Lys Lys Leu Val Leu Val Tyr Asp Gly Asp Lys Ala Gly  
 305 310 315 320  
 Gln Ala Ala Thr Leu Lys Ala Leu Asp Glu Ile Gly Asp Met Pro Val  
 325 330 335  
 Gln Ile Val Ser Met Pro Asp Asn Leu Asp Pro Asp Glu Tyr Leu Gln  
 340 345 350  
 40 Lys Asn Gly Pro Glu Asp Leu Ala Tyr Leu Leu Thr Lys Thr Arg Ile  
 355 360 365  
 Ser Pro Ile Glu Phe Tyr Ile His Gln Tyr Lys Pro Glu Asn Gly Glu  
 370 375 380  
 45 Asn Leu Gln Ala Gln Ile Glu Phe Leu Glu Lys Ile Ala Pro Leu Ile  
 385 390 395 400  
 Val Gln Glu Lys Ser Ile Ala Ala Gln Asn Ser Tyr Ile His Ile Leu  
 405 410 415  
 Ala Asp Ser Leu Ala Ser Phe Asp Tyr Thr Gln Ile Glu Gln Ile Val  
 420 425 430  
 55 Asn Glu Ser Arg Gln Val Gln Arg Gln Asn Arg Met Glu Arg Ile Ser  
 435 440 445  
 Arg Pro Thr Pro Ile Thr Met Pro Val Thr Lys Gln Leu Ser Ala Ile  
 450 455 460  
 60 Met Arg Ala Glu Ala His Leu Leu Tyr Arg Met Met Glu Ser Pro Leu

-160-

465                      470                      475                      480

Val Leu Asn Asp Tyr Arg Leu Arg Glu Asp Phe Ala Phe Ala Thr Pro  
                                  485                      490                      495

5      Glu Phe Gln Val Leu His Asp Leu Leu Gly Gln Tyr Gly Asn Leu Pro  
                                  500                      505                      510

10     Pro Glu Val Leu Ala Glu Gln Thr Glu Glu Val Glu Arg Ala Trp Tyr  
                                  515                      520                      525

       Gln Val Leu Ala Gln Asp Leu Pro Ala Glu Ile Ser Pro Gln Glu Leu  
                                  530                      535                      540

15     Ser Glu Val Glu Met Thr Arg Asn Lys Ala Leu Leu Asn Gln Asp Asn  
                                  545                      550                      555                      560

       Met Arg Ile Lys Lys Lys Val Gln Glu Ala Ser His Val Gly Asp Thr  
                                  565                      570                      575

20     Asp Thr Ala Leu Glu Glu Leu Glu Arg Leu Ile Ser Gln Lys Arg Arg  
                                  580                      585                      590

25     Met Glu

## (2) INFORMATION FOR SEQ ID NO:91:

30     (i) SEQUENCE CHARACTERISTICS:  
               (A) LENGTH: 900 base pairs  
               (B) TYPE: nucleic acid  
               (C) STRANDEDNESS: single  
               (D) TOPOLOGY: linear

35     (ii) MOLECULE TYPE: DNA (genomic)

       (iii) HYPOTHETICAL: NO

40     (iv) ANTI-SENSE: NO

       (ix) FEATURE:  
               (A) NAME/KEY: CDS  
               (B) LOCATION: 1..897  
 45         (D) OTHER INFORMATION: Era

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

50     ATG ACT TTT AAA TCA GGC TTT GTA GCC ATT TTA GGA CGT CCC AAT GTT                      48  
        Met Thr Phe Lys Ser Gly Phe Val Ala Ile Leu Gly Arg Pro Asn Val  
            1                                      5                                      10                                      15

55     GGG AAG TCA ACC TTT TTA AAT CAC GTT ATG GGG CAA AAG ATT GCC ATC                      96  
        Gly Lys Ser Thr Phe Leu Asn His Val Met Gly Gln Lys Ile Ala Ile  
                                  20                                      25                                      30

60     ATG AGT GAC AAG GCG CAG ACA ACG CGC AAT AAA ATC ATG GGA ATT TAC                      144  
        Met Ser Asp Thr Lys Ala Gln Thr Thr Arg Asn Lys Ile Met Gly Ile Tyr  
                                  35                                      40                                      45

-161-

	ACG ACT GAT AAG GAG CAA ATT GTC TTT ATC GAC ACA CCA GGG ATT CAC	192
	Thr Thr Asp Lys Glu Gln Ile Val Phe Ile Asp Thr Pro Gly Ile His	
	50 55 60	
5	AAA CCT AAA ACA GCT CTC GGA GAT TTC ATG GTT GAG TCT GCC TAC AGT	240
	Lys Pro Lys Thr Ala Leu Gly Asp Phe Met Val Glu Ser Ala Tyr Ser	
	65 70 75 80	
10	ACC CTT CGC GAA GTG GAC ACT GTT CTT TTC ATG GTG CCT GCT GAT GAA	288
	Thr Leu Arg Glu Val Asp Thr Val Leu Phe Met Val Pro Ala Asp Glu	
	85 90 95	
15	GCG CGT GGT AAG GGG GAC GAT ATG ATT ATC GAG CGT CTC AAG GCT GCC	336
	Ala Arg Gly Lys Gly Asp Asp Met Ile Ile Glu Arg Leu Lys Ala Ala	
	100 105 110	
20	AAG GTT CCT GTG ATT TTG GTG GTG AAT AAA ATC GAT AAG GTC CAT CCA	384
	Lys Val Pro Val Ile Leu Val Val Asn Lys Ile Asp Lys Val His Pro	
	115 120 125	
25	GAC CAG CTC TTG TCT CAG ATT GAT GAC TTC CGT AAT CAA ATG GAC TTT	432
	Asp Gln Leu Leu Ser Gln Ile Asp Asp Phe Arg Asn Gln Met Asp Phe	
	130 135 140	
30	AAG GAA ATT GTT CCA ATC TCA GCC CTT CAG GGA AAT AAC GTG TCT CGT	480
	Lys Glu Ile Val Pro Ile Ser Ala Leu Gln Gly Asn Asn Val Ser Arg	
	145 150 155 160	
35	CTA GTG GAT ATT TTG AGT GAA AAT CTG GAT GAA GGT TTC CAA TAT TTC	528
	Leu Val Asp Ile Leu Ser Glu Asn Leu Asp Glu Gly Phe Gln Tyr Phe	
	165 170 175	
40	CCG TCT GAT CAA ATC ACA GAT CAT CCA GAA CGT TTC TTA GTT TCA GAA	576
	Pro Ser Asp Gln Ile Thr Asp His Pro Glu Arg Phe Leu Val Ser Glu	
	180 185 190	
45	ATG GTT CGC GAG AAA GTC TTG CAC CTA ACT CGT GAA GAG ATT CCG CAT	624
	Met Val Arg Glu Lys Val Leu His Leu Thr Arg Glu Glu Ile Pro His	
	195 200 205	
50	TCT GTA GCA GTA GTT GTT GAC TCT ATG AAA CGA GAC GAA GAG ACA GAC	672
	Ser Val Ala Val Val Val Asp Ser Met Lys Arg Asp Glu Glu Thr Asp	
	210 215 220	
55	AAG GTT CAC ATC CGT GCA ACC ATC ATG GTC GAG CGC GAT AGC CAA AAA	720
	Lys Val His Ile Arg Ala Thr Ile Met Val Glu Arg Asp Ser Gln Lys	
	225 230 235 240	
60	GGG ATT ATC ATC GGT AAA GGT GGC GCT ATG CTT AAG AAA ATC GGT AGT	768
	Gly Ile Ile Ile Gly Lys Gly Gly Ala Met Leu Lys Lys Ile Gly Ser	
	245 250 255	
65	ATG GCC CGT CGT GAT ATC GAA CTC ATG CTA GGA GAC AAG GTC TTC CTA	816
	Met Ala Arg Arg Asp Ile Glu Leu Met Leu Gly Asp Lys Val Phe Leu	
	260 265 270	
70	GAA ACC TGG GTC AAG GTC AAG AAA AAC TGG CGC GAT AAA AAG CTA GAT	864
	Glu Thr Trp Val Lys Val Lys Lys Asn Trp Arg Asp Lys Lys Leu Asp	
	275 280 285	
75	TTG GCT GAC TTT GGC TAT AAT GAA AGA GAA TAC TAA	900

-162-

Leu Ala Asp Phe Gly Tyr Asn Glu Arg Glu Tyr  
 290 295

## 5 (2) INFORMATION FOR SEQ ID NO:92:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 299 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

15 Met Thr Phe Lys Ser Gly Phe Val Ala Ile Leu Gly Arg Pro Asn Val  
 1 5 10 15  
 20 Gly Lys Ser Thr Phe Leu Asn His Val Met Gly Gln Lys Ile Ala Ile  
 20 25 30  
 Met Ser Asp Lys Ala Gln Thr Thr Arg Asn Lys Ile Met Gly Ile Tyr  
 35 40 45  
 25 Thr Thr Asp Lys Glu Gln Ile Val Phe Ile Asp Thr Pro Gly Ile His  
 50 55 60  
 Lys Pro Lys Thr Ala Leu Gly Asp Phe Met Val Glu Ser Ala Tyr Ser  
 65 70 75 80  
 30 Thr Leu Arg Glu Val Asp Thr Val Leu Phe Met Val Pro Ala Asp Glu  
 85 90 95  
 Ala Arg Gly Lys Gly Asp Asp Met Ile Ile Glu Arg Leu Lys Ala Ala  
 100 105 110  
 Lys Val Pro Val Ile Leu Val Val Asn Lys Ile Asp Lys Val His Pro  
 115 120 125  
 40 Asp Gln Leu Leu Ser Gln Ile Asp Asp Phe Arg Asn Gln Met Asp Phe  
 130 135 140  
 Lys Glu Ile Val Pro Ile Ser Ala Leu Gln Gly Asn Asn Val Ser Arg  
 145 150 155 160  
 45 Leu Val Asp Ile Leu Ser Glu Asn Leu Asp Glu Gly Phe Gln Tyr Phe  
 165 170 175  
 Pro Ser Asp Gln Ile Thr Asp His Pro Glu Arg Phe Leu Val Ser Glu  
 180 185 190  
 Met Val Arg Glu Lys Val Leu His Leu Thr Arg Glu Glu Ile Pro His  
 195 200 205  
 55 Ser Val Ala Val Val Val Asp Ser Met Lys Arg Asp Glu Glu Thr Asp  
 210 215 220  
 Lys Val His Ile Arg Ala Thr Ile Met Val Glu Arg Asp Ser Gln Lys  
 225 230 235 240  
 60 Gly Ile Ile Ile Gly Lys Gly Gly Ala Met Leu Lys Lys Ile Gly Ser

-163-

		245		250		255										
	Met	Ala	Arg	Arg	Asp	Ile	Glu	Leu	Met	Leu	Gly	Asp	Lys	Val	Phe	Leu
			260						265					270		
5	Glu	Thr	Trp	Val	Lys	Val	Lys	Lys	Asn	Trp	Arg	Asp	Lys	Lys	Leu	Asp
			275					280					285			
10	Leu	Ala	Asp	Phe	Gly	Tyr	Asn	Glu	Arg	Glu	Tyr					
		290				295										

## (2) INFORMATION FOR SEQ ID NO:93:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1011 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1008

(D) OTHER INFORMATION: Gcp

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

35	ATG	AAG	GAT	AGA	TAT	ATT	TTA	GCA	TTT	GAG	ACA	TCC	TGT	GAT	GAG	ACC	48
	Met	Lys	Asp	Arg	Tyr	Ile	Leu	Ala	Phe	Glu	Thr	Ser	Cys	Asp	Glu	Thr	
	1				5					10					15		
40	AGT	GTC	GCC	GTC	TTG	AAA	AAC	GAC	GAT	GAG	CTC	TTG	TCC	AAT	GTC	ATT	96
	Ser	Val	Ala	Val	Leu	Lys	Asn	Asp	Asp	Glu	Leu	Leu	Ser	Asn	Val	Ile	
				20					25					30			
45	GCT	AGT	CAA	ATT	GAG	AGT	CAC	AAA	CGT	TTT	GGT	GGC	GTA	GTG	CCC	GAA	144
	Ala	Ser	Gln	Ile	Glu	Ser	His	Lys	Arg	Phe	Gly	Gly	Val	Val	Pro	Glu	
			35					40					45				
50	GTA	GCC	AGT	CGT	CAC	CAT	GTC	GAG	GTC	ATT	ACA	GCC	TGT	ATC	GAG	GAG	192
	Val	Ala	Ser	Arg	His	His	Val	Glu	Val	Ile	Thr	Ala	Cys	Ile	Glu	Glu	
		50					55					60					
55	GCA	TTG	GCA	GAA	GCA	GGG	ATT	ACC	GAA	GAG	GAC	GTG	ACA	GCT	GTT	GCG	240
	Ala	Leu	Ala	Glu	Ala	Gly	Ile	Thr	Glu	Glu	Asp	Val	Thr	Ala	Val	Ala	
		65				70				75					80		
60	GTT	ACC	TAC	GGA	CCA	GGC	TTG	GTC	GGA	GCC	TTG	CTA	GTT	GGT	TTG	TCA	288
	Val	Thr	Tyr	Gly	Pro	Gly	Leu	Val	Gly	Ala	Leu	Leu	Val	Gly	Leu	Ser	
				85			90							95			
60	GCT	GCC	AAG	GCC	TTT	GCT	TGG	GCT	CAC	GGA	CTT	CCA	CTG	ATT	CCT	GTT	336
	Ala	Ala	Lys	Ala	Phe	Ala	Trp	Ala	His	Gly	Leu	Pro	Leu	Ile	Pro	Val	



-164-

	100	105	110	
5	AAT CAC ATG GCT GGG CAC CTC ATG GCA GCT CAG AGT GTG GAG CCT TTG Asn His Met Ala Gly His Leu Met Ala Ala Gln Ser Val Glu Pro Leu 115 120 125	384		
10	GAG TTT CCC TTG CTA GCC CTT TTA GTC AGT GGT GGG CAC ACA GAG TTG Glu Phe Pro Leu Leu Ala Leu Leu Val Ser Gly Gly His Thr Glu Leu 130 135 140	432		
15	GTC TAT GTT TCT GAG GCT GGC GAT TAC AAG ATT GTT GGG GAG ACA CGA Val Tyr Val Ser Glu Ala Gly Asp Tyr Lys Ile Val Gly Glu Thr Arg 145 150 155 160	480		
20	GAC GAT GCA GTT GGG GAG GCT TAT GAC AAG GTC GGT CGT GTC ATG GGC Asp Asp Ala Val Gly Glu Ala Tyr Asp Lys Val Gly Arg Val Met Gly 165 170 175	528		
25	TTG ACC TAT CCT GCA GGT CGT GAG ATT GAC GAG CTG GCT CAT CAG GGG Leu Thr Tyr Pro Ala Gly Arg Glu Ile Asp Glu Leu Ala His Gln Gly 180 185 190	576		
30	CAC GAT ATT TAT GAT TTC CCC CGT GCC ATG ATT AAG GAA GAT AAT CTG His Asp Ile Tyr Asp Phe Pro Arg Ala Met Ile Lys Glu Asp Asn Leu 195 200 205	624		
35	GAG TTC TCC TTC TCA GGT TTG AAA TCT GCC TTT ATC AAT CTT CAT CAC Glu Phe Ser Phe Ser Gly Leu Lys Ser Ala Phe Ile Asn Leu His His 210 215 220	672		
40	AAT GCC GAG CAA AAG GGA GAA AGC CTG TCT ACA GAA GAT TTG TGT GCT Asn Ala Glu Gln Lys Gly Glu Ser Leu Ser Thr Glu Asp Leu Cys Ala 225 230 235 240	720		
45	TCC TTC CAA GCA GCA GTT ATG GAC ATT CTC ATG GCA AAA ACC AAG AAG Ser Phe Gln Ala Ala Val Met Asp Ile Leu Met Ala Lys Thr Lys Lys 245 250 255	768		
50	GCT TTG GAG AAA TAT CCT GTT AAA ACC CTA GTT GTG GCA GGT GGT GTG Ala Leu Glu Lys Tyr Pro Val Lys Thr Leu Val Val Ala Gly Gly Val 260 265 270	816		
55	GCA GCC AAT AAA GGT CTC AGA GAA CGC CTA GCA ACT GAA ATC ACA GAT Ala Ala Asn Lys Gly Leu Arg Glu Arg Leu Ala Thr Glu Ile Thr Asp 275 280 285	864		
60	GTC AAT GTT ATC ATT CCA CCT CTG CGT CTC TGC GGA GAC AAT GCA GGT Val Asn Val Ile Ile Pro Pro Leu Arg Leu Cys Gly Asp Asn Ala Gly 290 295 300	912		
65	ATG ATT GCT TAT GCC AGT GTC AGC GAG TGG AAC AAA GAA AAC TTT GCA Met Ile Ala Tyr Ala Ser Val Ser Glu Trp Asn Lys Glu Asn Phe Ala 305 310 315 320	960		
70	AAC TTG GAC CTC AAT GCC AAA CCA AGT CTT GCC TTT GAT ACC ATG GAA Asn Leu Asp Leu Asn Ala Lys Pro Ser Leu Ala Phe Asp Thr Met Glu 325 330 335	1008		
75	TAA	1011		

-165-

## (2) INFORMATION FOR SEQ ID NO:94:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 336 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

	Met	Lys	Asp	Arg	Tyr	Ile	Leu	Ala	Phe	Glu	Thr	Ser	Cys	Asp	Glu	Thr	
	1				5					10					15		
15	Ser	Val	Ala	Val	Leu	Lys	Asn	Asp	Asp	Glu	Leu	Leu	Ser	Asn	Val	Ile	
				20					25						30		
	Ala	Ser	Gln	Ile	Glu	Ser	His	Lys	Arg	Phe	Gly	Gly	Val	Val	Pro	Glu	
			35					40					45				
20	Val	Ala	Ser	Arg	His	His	Val	Glu	Val	Ile	Thr	Ala	Cys	Ile	Glu	Glu	
		50					55					60					
	Ala	Leu	Ala	Glu	Ala	Gly	Ile	Thr	Glu	Glu	Asp	Val	Thr	Ala	Val	Ala	
25		65				70				75						80	
	Val	Thr	Tyr	Gly	Pro	Gly	Leu	Val	Gly	Ala	Leu	Leu	Val	Gly	Leu	Ser	
					85					90					95		
30	Ala	Ala	Lys	Ala	Phe	Ala	Trp	Ala	His	Gly	Leu	Pro	Leu	Ile	Pro	Val	
				100					105						110		
	Asn	His	Met	Ala	Gly	His	Leu	Met	Ala	Ala	Gln	Ser	Val	Glu	Pro	Leu	
			115					120						125			
35	Glu	Phe	Pro	Leu	Leu	Ala	Leu	Leu	Val	Ser	Gly	Gly	His	Thr	Glu	Leu	
		130				135						140					
	Val	Tyr	Val	Ser	Glu	Ala	Gly	Asp	Tyr	Lys	Ile	Val	Gly	Glu	Thr	Arg	
40		145				150					155					160	
	Asp	Asp	Ala	Val	Gly	Glu	Ala	Tyr	Asp	Lys	Val	Gly	Arg	Val	Met	Gly	
				165					170						175		
45	Leu	Thr	Tyr	Pro	Ala	Gly	Arg	Glu	Ile	Asp	Glu	Leu	Ala	His	Gln	Gly	
				180					185						190		
	His	Asp	Ile	Tyr	Asp	Phe	Pro	Arg	Ala	Met	Ile	Lys	Glu	Asp	Asn	Leu	
			195					200					205				
50	Glu	Phe	Ser	Phe	Ser	Gly	Leu	Lys	Ser	Ala	Phe	Ile	Asn	Leu	His	His	
		210				215						220					
	Asn	Ala	Glu	Gln	Lys	Gly	Glu	Ser	Leu	Ser	Thr	Glu	Asp	Leu	Cys	Ala	
55		225				230					235					240	
	Ser	Phe	Gln	Ala	Ala	Val	Met	Asp	Ile	Leu	Met	Ala	Lys	Thr	Lys	Lys	
				245					250						255		
60	Ala	Leu	Glu	Lys	Tyr	Pro	Val	Lys	Thr	Leu	Val	Val	Ala	Gly	Gly	Val	
			260						265					270			

-166-

Ala Ala Asn Lys Gly Leu Arg Glu Arg Leu Ala Thr Glu Ile Thr Asp  
 275 280 285

5 Val Asn Val Ile Ile Pro Pro Leu Arg Leu Cys Gly Asp Asn Ala Gly  
 290 295 300

Met Ile Ala Tyr Ala Ser Val Ser Glu Trp Asn Lys Glu Asn Phe Ala  
 305 310 315 320

10 Asn Leu Asp Leu Asn Ala Lys Pro Ser Leu Ala Phe Asp Thr Met Glu  
 325 330 335

## (2) INFORMATION FOR SEQ ID NO:95:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 774 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(ix) FEATURE:  
 30 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..771  
 (D) OTHER INFORMATION: HI0454

## 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

ATG ATT TTT GAT ACA CAT ACA CAC TTG AAT GTA GAA GAA TTT GCA GGT	48
Met Ile Phe Asp Thr His Thr His Leu Asn Val Glu Glu Phe Ala Gly	
1 5 10 15	
CGT GAG GCA GAA GAA ATT GCC TTG GCT GCT GAG ATG GGT GTG ACA CAG	96
Arg Glu Ala Glu Glu Ile Ala Leu Ala Ala Glu Met Gly Val Thr Gln	
20 25 30	
ATG AAT ATT GTT GGT TTT GAT AAA CCG ACG ATT GAG CAT GCC TTG GAG	144
Met Asn Ile Val Gly Phe Asp Lys Pro Thr Ile Glu His Ala Leu Glu	
35 40 45	
TTG GTA GAT GAG TAT GAG CAG CTC TAT GCG ACT ATT GGT TGG CAT CCT	192
Leu Val Asp Glu Tyr Glu Gln Leu Tyr Ala Thr Ile Gly Trp His Pro	
50 55 60	
ACA GAA GCT GGT ACT TAT ACA GAG GAA GTT GAG GCT TAC TTG TTG GAT	240
Thr Glu Ala Gly Thr Tyr Thr Glu Glu Val Glu Ala Tyr Leu Leu Asp	
55 65 70 75 80	
AAG TTA AAA CAT TCC AAG GTT GTG GCT TTA GGT GAA ATT GGC TTA GAC	288
Lys Leu Lys His Ser Lys Val Val Ala Leu Gly Glu Ile Gly Leu Asp	
85 90 95	
TAC CAT TGG ATG ACA GCG CCC AAA GAG GTG CAG GAG CAG GTT TTT CGC	336

-167-

	Tyr	His	Trp	Met	Thr	Ala	Pro	Lys	Glu	Val	Gln	Glu	Gln	Val	Phe	Arg	
				100					105					110			
5	CGT	CAG	ATT	CAG	CTA	TCT	AAG	GAC	TTG	GAT	TTG	CCT	TTT	GTT	GTC	CAT	384
	Arg	Gln	Ile	Gln	Leu	Ser	Lys	Asp	Leu	Asp	Leu	Pro	Phe	Val	Val	His	
			115					120					125				
10	ACC	CGT	GAT	GCG	CTG	GAA	GAT	ACC	TAT	GAG	ATT	ATC	AAG	AGT	GAG	GGC	432
	Thr	Arg	Asp	Ala	Leu	Glu	Asp	Thr	Tyr	Glu	Ile	Ile	Lys	Ser	Glu	Gly	
		130					135				140						
15	GTT	GGT	CCT	CGT	GGT	GGT	ATC	ATG	CAT	TCA	TTT	TCA	GGG	ACG	CTT	GAG	480
	Val	Gly	Pro	Arg	Gly	Gly	Ile	Met	His	Ser	Phe	Ser	Gly	Thr	Leu	Glu	
	145					150				155						160	
20	TGG	GCA	GAG	AAG	TTT	GTG	GAT	CTT	GGT	ATG	ACC	ATT	TCC	TTC	TCA	GGA	528
	Trp	Ala	Glu	Lys	Phe	Val	Asp	Leu	Gly	Met	Thr	Ile	Ser	Phe	Ser	Gly	
				165					170						175		
25	GTG	GTG	ACC	TTC	AAG	AAG	GCA	ACT	GAC	CTC	CAA	GAA	GCA	GCT	AAA	GAG	576
	Val	Val	Thr	Phe	Lys	Lys	Ala	Thr	Asp	Leu	Gln	Glu	Ala	Ala	Lys	Glu	
				180					185					190			
30	TTA	CCT	TTG	GAC	AAG	ATG	TTG	GTA	GAA	ACA	GAT	GCG	CCT	TAC	TTA	GCA	624
	Leu	Pro	Leu	Asp	Lys	Met	Leu	Val	Glu	Thr	Asp	Ala	Pro	Tyr	Leu	Ala	
			195				200						205				
35	CCT	GTA	CCC	AAG	CGT	GGT	CGT	GAA	AAT	AAA	ACA	GCC	TAT	ACT	CGC	TAT	672
	Pro	Val	Pro	Lys	Arg	Gly	Arg	Glu	Asn	Lys	Thr	Ala	Tyr	Thr	Arg	Tyr	
		210				215						220					
40	GTG	GTC	GAC	TTT	ATC	GCT	GAC	TTG	CGT	GGT	ATG	ACG	ACA	GAA	GAG	CTG	720
	Val	Val	Asp	Phe	Ile	Ala	Asp	Leu	Arg	Gly	Met	Thr	Thr	Glu	Glu	Leu	
	225				230					235					240		
45	GCG	GTA	GCA	ACG	ACT	GCA	AAT	GCA	GAA	CGC	ATT	TTT	GGA	TTG	GAC	AGC	768
	Ala	Val	Ala	Thr	Thr	Ala	Asn	Ala	Glu	Arg	Ile	Phe	Gly	Leu	Asp	Ser	
				245				250							255		
50	AAG	TAA															774
	Lys																

## 45 (2) INFORMATION FOR SEQ ID NO:96:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 257 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

55	Met	Ile	Phe	Asp	Thr	His	Thr	His	Leu	Asn	Val	Glu	Glu	Phe	Ala	Gly
	1				5					10					15	
60	Arg	Glu	Ala	Glu	Ile	Ala	Leu	Ala	Ala	Glu	Met	Gly	Val	Thr	Gln	
			20				25						30			

-168-

Met Asn Ile Val Gly Phe Asp Lys Pro Thr Ile Glu His Ala Leu Glu  
                   35                                  40                                  45

5 Leu Val Asp Glu Tyr Glu Gln Leu Tyr Ala Thr Ile Gly Trp His Pro  
                   50                                  55                                  60

Thr Glu Ala Gly Thr Tyr Thr Glu Glu Val Glu Ala Tyr Leu Leu Asp  
                   65                                  70                                  75                                  80

10 Lys Leu Lys His Ser Lys Val Val Ala Leu Gly Glu Ile Gly Leu Asp  
                                   85                                  90                                  95

Tyr His Trp Met Thr Ala Pro Lys Glu Val Gln Glu Gln Val Phe Arg  
                                   100                                  105                                  110

15 Arg Gln Ile Gln Leu Ser Lys Asp Leu Asp Leu Pro Phe Val Val His  
                                   115                                  120                                  125

20 Thr Arg Asp Ala Leu Glu Asp Thr Tyr Glu Ile Ile Lys Ser Glu Gly  
                   130                                  135                                  140

Val Gly Pro Arg Gly Gly Ile Met His Ser Phe Ser Gly Thr Leu Glu  
                   145                                  150                                  155                                  160

25 Trp Ala Glu Lys Phe Val Asp Leu Gly Met Thr Ile Ser Phe Ser Gly  
                                   165                                  170                                  175

Val Val Thr Phe Lys Lys Ala Thr Asp Leu Gln Glu Ala Ala Lys Glu  
                                   180                                  185                                  190

30 Leu Pro Leu Asp Lys Met Leu Val Glu Thr Asp Ala Pro Tyr Leu Ala  
                                   195                                  200                                  205

35 Pro Val Pro Lys Arg Gly Arg Glu Asn Lys Thr Ala Tyr Thr Arg Tyr  
                   210                                  215                                  220

Val Val Asp Phe Ile Ala Asp Leu Arg Gly Met Thr Thr Glu Glu Leu  
                   225                                  230                                  235                                  240

40 Ala Val Ala Thr Thr Ala Asn Ala Glu Arg Ile Phe Gly Leu Asp Ser  
                                   245                                  250                                  255

Lys

45

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1959 base pairs  
 50 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)  
 55

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

60

(ix) FEATURE:

-169-

(A) NAME/KEY: CDS  
 (B) LOCATION: 1..1959  
 (D) OTHER INFORMATION: Ligase

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

10	ATG AAT AAA AGA ATG AAT GAG TTA GTC GCT TTG CTC AAT CGC TAT GCG Met Asn Lys Arg Met Asn Glu Leu Val Ala Leu Leu Asn Arg Tyr Ala	48
	1 5 10 15	
15	ACT GAG TAC TAT ACC AGC GAT AAT CCC TCG GTT TCA GAC AGT GAG TAT Thr Glu Tyr Tyr Thr Ser Asp Asn Pro Ser Val Ser Asp Ser Glu Tyr	96
	20 25 30	
20	GAC CGC CTT TAC CGT GAG TTG GTC GAG TTA GAA ACT GCT TAT CCA GAG Asp Arg Leu Tyr Arg Glu Leu Val Glu Leu Glu Thr Ala Tyr Pro Glu	144
	35 40 45	
25	CAA GTG CTA GCA GAC AGT CCG ACT CAT CGT GTT GGT GGC AAG GTT TTA Gln Val Leu Ala Asp Ser Pro Thr His Arg Val Gly Gly Lys Val Leu	192
	50 55 60	
30	GAT GGT TTT GAA AAA TAC AGT CAT CAG TAT CCT CTT TAT AGT TTG CAG Asp Gly Phe Glu Lys Tyr Ser His Gln Tyr Pro Leu Tyr Ser Leu Gln	240
	65 70 75 80	
35	GAT GCT TTT TCA CGT GAG GAG CTA GAT GCT TTT GAT GCG CGT GTT CGT Asp Ala Phe Ser Arg Glu Glu Leu Asp Ala Phe Asp Ala Arg Val Arg	288
	85 90 95	
40	AAG GAA GTG GCT CAT CCG ACC TAT ATT TGT GAG CTG AAA ATC GAT GGC Lys Glu Val Ala His Pro Thr Tyr Ile Cys Glu Leu Lys Ile Asp Gly	336
	100 105 110	
45	TTA TCT ATC TCG CTG ACT TAT GAA AAG GGG ATT TTG GTT GCT GGG GTA Leu Ser Ile Ser Leu Thr Tyr Glu Lys Gly Ile Leu Val Ala Gly Val	384
	115 120 125	
50	ACA CGT GGA GAT GGT TCA ATT GGT GAA AAT ATC ACA GAA AAC CTC AAG Thr Arg Gly Asp Gly Ser Ile Gly Glu Asn Ile Thr Glu Asn Leu Lys	432
	130 135 140	
55	CGT GTT AAG GAC ATC CCT TTG ACT TTG CCA GAA GAA CTA GAT ATC ACA Arg Val Lys Asp Ile Pro Leu Thr Leu Pro Glu Glu Leu Asp Ile Thr	480
	145 150 155 160	
60	GTT CGT GGG GAA TGT TAC ATG CCA CGC GCT TCC TTT GAC CAA GTT AAC Val Arg Gly Glu Cys Tyr Met Pro Arg Ala Ser Phe Asp Gln Val Asn	528
	165 170 175	
65	CAA GCG CGC CAA GAA AAT GGA GAG CCT GAA TTT GCT AAT CCT CGT AAT Gln Ala Arg Gln Glu Asn Gly Glu Pro Glu Phe Ala Asn Pro Arg Asn	576
	180 185 190	
70	GCG GCA GCA GGA ACT CTG CGT CAG TTG GAT ACA GCA GTA GTT GCC AAG Ala Ala Ala Gly Thr Leu Arg Gln Leu Asp Thr Ala Val Val Ala Lys	624
	195 200 205	
75	CGT AAT CTT GCA ACG TTT CTC TAT CAA GAA GCC AGC CCT TCA ACT CGT Arg Asn Leu Ala Thr Phe Leu Tyr Gln Glu Ala Ser Pro Ser Thr Arg	672

-170-

	210	215	220	
5	GAT AGC CAA GAA AAG GGT TTG AAG TAC CTA GAA CAA CTA GGT TTT GTG Asp Ser Gln Glu Lys Gly Leu Lys Tyr Leu Glu Gln Leu Gly Phe Val 225 230 235 240			720
10	GTC AAT CCT AAG CGA ATC TTG GCT GAA AAC ATA GAT GAA ATC TGG AAT Val Asn Pro Lys Arg Ile Leu Ala Glu Asn Ile Asp Glu Ile Trp Asn 245 250 255			768
15	TTT ATC CAA GAA GTA GGA CAG GAA CGG GAA AAT CTG CCT TAC GAT ATT Phe Ile Gln Glu Val Gly Gln Glu Arg Glu Asn Leu Pro Tyr Asp Ile 260 265 270			816
20	GAT GGA GTG GTA ATC AAG GTC AAC GAC CTA GCA AGT CAA GAA GAA CTT Asp Gly Val Val Ile Lys Val Asn Asp Leu Ala Ser Gln Glu Glu Leu 275 280 285			864
25	GGT TTT ACC GTT AAG GCT CCA AAG TGG GCA GTA GCC TAC AAG TTC CCT Gly Phe Thr Val Lys Ala Pro Lys Trp Ala Val Ala Tyr Lys Phe Pro 290 295 300			912
30	GCT GAA GAA AAA GAA GCT CAA CTC TTA TCA GTT GAC TGG ACA GTT GGC Ala Glu Glu Lys Glu Ala Gln Leu Leu Ser Val Asp Trp Thr Val Gly 305 310 315 320			960
35	CGT ACC GGT GTT GTA ACT CCA ACT GCT AAT CTA ACA CCA GTA CAA CTT Arg Thr Gly Val Val Thr Pro Thr Ala Asn Leu Thr Pro Val Gln Leu 325 330 335			1008
40	GCC GGT ACG ACT GTT AGC CGT GCG ACC CTG CAC AAT GTA GAT TAT ATT Ala Gly Thr Thr Val Ser Arg Ala Thr Leu His Asn Val Asp Tyr Ile 340 345 350			1056
45	GCT GAA AAA GAT ATC CGA AAA GAC GAT ACG GTC ATT GTA TAT AAG GCT Ala Glu Lys Asp Ile Arg Lys Asp Asp Thr Val Ile Val Tyr Lys Ala 355 360 365			1104
50	GGT GAC ATC ATC CCT GCC GTT TTA CGT GTG GTA GAG TCC AAA CGG GTT Gly Asp Ile Ile Pro Ala Val Leu Arg Val Val Glu Ser Lys Arg Val 370 375 380			1152
55	TCT GAA GAA AAA CTA GAT ATC CCT ACA AAC TGT CCA AGT TGT AAC TCT Ser Glu Glu Lys Leu Asp Ile Pro Thr Asn Cys Pro Ser Cys Asn Ser 385 390 395 400			1200
60	GAC TTG TTG CAC TTT GAA GAT GAA GTG GCC CTA CGT TGT ATC AAT CCG Asp Leu Leu His Phe Glu Asp Glu Val Ala Leu Arg Cys Ile Asn Pro 405 410 415			1248
65	CGT TGC CCT GCT CAA ATC ATG GAA GGC TTG ATT CAC TTT GCT TCT CGT Arg Cys Pro Ala Gln Ile Met Glu Gly Leu Ile His Phe Ala Ser Arg 420 425 430			1296
70	GAT GCT ATG AAT ATT ACA GGC CTT GGT CCA TCT ATT GTT GAG AAG CTT Asp Ala Met Asn Ile Thr Gly Leu Gly Pro Ser Ile Val Glu Lys Leu 435 440 445			1344
75	TTT GCT GCT AAT TTA GTC AAG GAT GTG GCG GAT ATT TAT CGT TTG CAA Phe Ala Ala Asn Leu Val Lys Asp Val Ala Asp Ile Tyr Arg Leu Gln 450 455 460			1392

-171-

	GAA GAG GAT TTC CTC CTT TTA GAG GGG GTT AAG GAA AAG TCC GCT GCT	1440
	Glu Glu Asp Phe Leu Leu Leu Glu Gly Val Lys Glu Lys Ser Ala Ala	
	465 470 475 480	
5	AAA CTG TAT CAG GCT ATC CAA GCA TCA AAG GAA AAT TCT GCC GAG AAG	1488
	Lys Leu Tyr Gln Ala Ile Gln Ala Ser Lys Glu Asn Ser Ala Glu Lys	
	485 490 495	
10	CTC TTA TTT GGT TTG GGA ATT CGT CAT GTC GGA AGC AAG GCT AGT CAG	1536
	Leu Leu Phe Gly Leu Gly Ile Arg His Val Gly Ser Lys Ala Ser Gln	
	500 505 510	
15	CTT TTA CTT CAA TAT TTC CAT TCA ATT GAA AAT CTG TAT CAG GCA GAT	1584
	Leu Leu Leu Gln Tyr Phe His Ser Ile Glu Asn Leu Tyr Gln Ala Asp	
	515 520 525	
20	TCA GAG GAA GTG GCT AGT ATT GAA AGT CTA GGT GGC GTG ATT GCC AAA	1632
	Ser Glu Glu Val Ala Ser Ile Glu Ser Leu Gly Gly Val Ile Ala Lys	
	530 535 540	
25	AGT CTT CAG ACT TAT TTT GCG GCA GAA GGC TCT GAA ATT CTG CTC AGA	1680
	Ser Leu Gln Thr Tyr Phe Ala Ala Glu Gly Ser Glu Ile Leu Leu Arg	
	545 550 555 560	
	GAA TTG AAA GAA ACT GGG GTC AAT CTG GAC TAT AAA GGA CAG ACG GTA	1728
	Glu Leu Lys Glu Thr Gly Val Asn Leu Asp Tyr Lys Gly Gln Thr Val	
	565 570 575	
30	GTA GCG GAT GCG GCC TTG TCA GGT TTG ACC GTG GTA TTG ACA GGA AAA	1776
	Val Ala Asp Ala Ala Leu Ser Gly Leu Thr Val Val Leu Thr Gly Lys	
	580 585 590	
35	TTG GAA CGA CTC AAG CGC TCA GAA GCT AAA AGT AAA CTC GAA AGT CTG	1824
	Leu Glu Arg Leu Lys Arg Ser Glu Ala Lys Ser Lys Leu Glu Ser Leu	
	595 600 605	
40	GGT GCC AAA GTG ACA GGT AGT GTT TCT AAA AAG ACC GAC CTC GTC GTG	1872
	Gly Ala Lys Val Thr Gly Ser Val Ser Lys Lys Thr Asp Leu Val Val	
	610 615 620	
	GTA GGT GCA GAC GCT GGA AGT AAA CTG CAA AAA GCA CAA GAA CTT GGT	1920
	Val Gly Ala Asp Ala Gly Ser Lys Leu Gln Lys Ala Gln Glu Leu Gly	
	625 630 635 640	
45	ATC CAG GTC AGA GAT GAG GCA TGG CTA GAA AGT TTG TAA	1959
	Ile Gln Val Arg Asp Glu Ala Trp Leu Glu Ser Leu	
	645 650	

50

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

55

(A) LENGTH: 653 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:



-172-

	Met	Asn	Lys	Arg	Met	Asn	Glu	Leu	Val	Ala	Leu	Leu	Asn	Arg	Tyr	Ala	
	1				5					10					15		
5	Thr	Glu	Tyr	Tyr	Thr	Ser	Asp	Asn	Pro	Ser	Val	Ser	Asp	Ser	Glu	Tyr	
				20					25					30			
	Asp	Arg	Leu	Tyr	Arg	Glu	Leu	Val	Glu	Leu	Glu	Thr	Ala	Tyr	Pro	Glu	
			35					40					45				
10	Gln	Val	Leu	Ala	Asp	Ser	Pro	Thr	His	Arg	Val	Gly	Gly	Lys	Val	Leu	
		50					55					60					
	Asp	Gly	Phe	Glu	Lys	Tyr	Ser	His	Gln	Tyr	Pro	Leu	Tyr	Ser	Leu	Gln	
	65				70					75						80	
15	Asp	Ala	Phe	Ser	Arg	Glu	Glu	Leu	Asp	Ala	Phe	Asp	Ala	Arg	Val	Arg	
					85					90					95		
20	Lys	Glu	Val	Ala	His	Pro	Thr	Tyr	Ile	Cys	Glu	Leu	Lys	Ile	Asp	Gly	
				100					105					110			
	Leu	Ser	Ile	Ser	Leu	Thr	Tyr	Glu	Lys	Gly	Ile	Leu	Val	Ala	Gly	Val	
			115					120					125				
25	Thr	Arg	Gly	Asp	Gly	Ser	Ile	Gly	Glu	Asn	Ile	Thr	Glu	Asn	Leu	Lys	
		130					135					140					
	Arg	Val	Lys	Asp	Ile	Pro	Leu	Thr	Leu	Pro	Glu	Glu	Leu	Asp	Ile	Thr	
	145				150						155				160		
30	Val	Arg	Gly	Glu	Cys	Tyr	Met	Pro	Arg	Ala	Ser	Phe	Asp	Gln	Val	Asn	
					165				170						175		
	Gln	Ala	Arg	Gln	Glu	Asn	Gly	Glu	Pro	Glu	Phe	Ala	Asn	Pro	Arg	Asn	
			180					185					190				
35	Ala	Ala	Ala	Gly	Thr	Leu	Arg	Gln	Leu	Asp	Thr	Ala	Val	Val	Ala	Lys	
			195					200					205				
40	Arg	Asn	Leu	Ala	Thr	Phe	Leu	Tyr	Gln	Glu	Ala	Ser	Pro	Ser	Thr	Arg	
		210				215						220					
	Asp	Ser	Gln	Glu	Lys	Gly	Leu	Lys	Tyr	Leu	Glu	Gln	Leu	Gly	Phe	Val	
	225				230						235				240		
45	Val	Asn	Pro	Lys	Arg	Ile	Leu	Ala	Glu	Asn	Ile	Asp	Glu	Ile	Trp	Asn	
				245					250					255			
	Phe	Ile	Gln	Glu	Val	Gly	Gln	Glu	Arg	Glu	Asn	Leu	Pro	Tyr	Asp	Ile	
			260					265						270			
50	Asp	Gly	Val	Val	Ile	Lys	Val	Asn	Asp	Leu	Ala	Ser	Gln	Glu	Glu	Leu	
			275					280					285				
55	Gly	Phe	Thr	Val	Lys	Ala	Pro	Lys	Trp	Ala	Val	Ala	Tyr	Lys	Phe	Pro	
		290				295						300					
	Ala	Glu	Glu	Lys	Glu	Ala	Gln	Leu	Leu	Ser	Val	Asp	Trp	Thr	Val	Gly	
	305				310					315					320		
60	Arg	Thr	Gly	Val	Val	Thr	Pro	Thr	Ala	Asn	Leu	Thr	Pro	Val	Gln	Leu	

-173-

	325	330	335
	Ala Gly Thr Thr Val Ser Arg Ala Thr Leu His Asn Val Asp Tyr Ile		
	340	345	350
5	Ala Glu Lys Asp Ile Arg Lys Asp Asp Thr Val Ile Val Tyr Lys Ala		
	355	360	365
10	Gly Asp Ile Ile Pro Ala Val Leu Arg Val Val Glu Ser Lys Arg Val		
	370	375	380
	Ser Glu Glu Lys Leu Asp Ile Pro Thr Asn Cys Pro Ser Cys Asn Ser		
	385	390	395
15	Asp Leu Leu His Phe Glu Asp Glu Val Ala Leu Arg Cys Ile Asn Pro		
	405	410	415
	Arg Cys Pro Ala Gln Ile Met Glu Gly Leu Ile His Phe Ala Ser Arg		
	420	425	430
20	Asp Ala Met Asn Ile Thr Gly Leu Gly Pro Ser Ile Val Glu Lys Leu		
	435	440	445
	Phe Ala Ala Asn Leu Val Lys Asp Val Ala Asp Ile Tyr Arg Leu Gln		
25	450	455	460
	Glu Glu Asp Phe Leu Leu Leu Glu Gly Val Lys Glu Lys Ser Ala Ala		
	465	470	475
30	Lys Leu Tyr Gln Ala Ile Gln Ala Ser Lys Glu Asn Ser Ala Glu Lys		
	485	490	495
	Leu Leu Phe Gly Leu Gly Ile Arg His Val Gly Ser Lys Ala Ser Gln		
	500	505	510
35	Leu Leu Leu Gln Tyr Phe His Ser Ile Glu Asn Leu Tyr Gln Ala Asp		
	515	520	525
	Ser Glu Glu Val Ala Ser Ile Glu Ser Leu Gly Gly Val Ile Ala Lys		
40	530	535	540
	Ser Leu Gln Thr Tyr Phe Ala Ala Glu Gly Ser Glu Ile Leu Leu Arg		
	545	550	555
45	Glu Leu Lys Glu Thr Gly Val Asn Leu Asp Tyr Lys Gly Gln Thr Val		
	565	570	575
	Val Ala Asp Ala Ala Leu Ser Gly Leu Thr Val Val Leu Thr Gly Lys		
	580	585	590
50	Leu Glu Arg Leu Lys Arg Ser Glu Ala Lys Ser Lys Leu Glu Ser Leu		
	595	600	605
	Gly Ala Lys Val Thr Gly Ser Val Ser Lys Lys Thr Asp Leu Val Val		
55	610	615	620
	Val Gly Ala Asp Ala Gly Ser Lys Leu Gln Lys Ala Gln Glu Leu Gly		
	625	630	635
60	Ile Gln Val Arg Asp Glu Ala Trp Leu Glu Ser Leu		
	645	650	

-174-

## (2) INFORMATION FOR SEQ ID NO:99:

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 981 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..981  
 (D) OTHER INFORMATION: MraY

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

25	ATG TTT ATT TCC ATC AGT GCT GGA ATT GTG ACA TTT TTA CTA ACT TTA	48
	Met Phe Ile Ser Ile Ser Ala Gly Ile Val Thr Phe Leu Leu Thr Leu	
	1 5 10 15	
30	GTA GGA ATT CCG GCC TTT ATC CAA TTT TAT AGA AAG GCG CAA ATT ACA	96
	Val Gly Ile Pro Ala Phe Ile Gln Phe Tyr Arg Lys Ala Gln Ile Thr	
	20 25 30	
35	GGC CAG CAG ATG CAT GAG GAT GTC AAA CAG CAT CAG GCA AAA GCT GGG	144
	Gly Gln Gln Met His Glu Asp Val Lys Gln His Gln Ala Lys Ala Gly	
	35 40 45	
40	ACT CCT ACA ATG GGA GGT TTG GTT TTC TTG ATT ACT TCT GTT TTG GTT	192
	Thr Pro Thr Met Gly Gly Leu Val Phe Leu Ile Thr Ser Val Leu Val	
	50 55 60	
45	GCT TTC TTT TTC GCC CTA TTT AGT AGC CAA TTC AGC AAT AAT GTG GGA	240
	Ala Phe Phe Phe Ala Leu Phe Ser Ser Gln Phe Ser Asn Asn Val Gly	
	65 70 75 80	
50	ATG ATT TTG TTC ATC TTG GTC TTG TAT GGC TTG GTC GGA TTT TTA GAT	288
	Met Ile Leu Phe Ile Leu Val Leu Tyr Gly Leu Val Gly Phe Leu Asp	
	85 90 95	
55	GAC TTT CTC AAG GTC TTT CGT AAA ATC AAT GAG GGG CTT AAT CCT AAG	336
	Asp Phe Leu Lys Val Phe Arg Lys Ile Asn Glu Gly Leu Asn Pro Lys	
	100 105 110	
60	CAA AAA TTA GCT CTT CAG CTT CTA GGT GGA GTT ATC TTC TAT CTT TTC	384
	Gln Lys Leu Ala Leu Gln Leu Leu Gly Gly Val Ile Phe Tyr Leu Phe	
	115 120 125	
65	TAT GAG CGC GGT GGC GAT ATC CTG TCT GTC TTT GGT TAT CCA GTT CAT	432
	Tyr Glu Arg Gly Gly Asp Ile Leu Ser Val Phe Gly Tyr Pro Val His	
	130 135 140	
70	TTG GGA TTT TTC TAT ATT TTC TTC GCT CTT TTC TGG CTA GTC GGT TTT	480
	Leu Gly Phe Phe Tyr Ile Phe Phe Ala Leu Phe Trp Leu Val Gly Phe	

-175-

	145		150		155		160	
	TCA AAC GCA GTA AAC TTG ACA GAC GGT GTT GAC GGT TTA GCT AGT ATT							528
5	Ser Asn Ala Val Asn Leu Thr Asp Gly Val Asp Gly Leu Ala Ser Ile	165		170		175		
	TCC GTT GTG ATT AGT TTG TTT GCC TAT GGA GTT ATT GCC TAT GTG CAA							576
	Ser Val Val Ile Ser Leu Phe Ala Tyr Gly Val Ile Ala Tyr Val Gln	180		185		190		
10	GGT CAG ATG GAT ATT CTT CTA GTG ATT CTT GCC ATG ATT GGT GGT TTG							624
	Gly Gln Met Asp Ile Leu Leu Val Ile Leu Ala Met Ile Gly Gly Leu	195		200		205		
15	CTC GGT TTC TTC ATC TTT AAC CAT AAG CCT GCC AAG GTC TTT ATG GGT							672
	Leu Gly Phe Phe Ile Phe Asn His Lys Pro Ala Lys Val Phe Met Gly	210		215		220		
20	GAT GTG GGA AGT TTG GCC CTA GGT GGG ATG CTG GCA GCT ATC TCT ATG							720
	Asp Val Gly Ser Leu Ala Leu Gly Gly Met Leu Ala Ala Ile Ser Met	225		230		235		240
	GCT CTC CAC CAG GAA TGG ACT CTC TTG ATT ATC GGA ATT GTG TAT GTT							768
25	Ala Leu His Gln Trp Thr Leu Leu Ile Ile Gly Ile Val Tyr Val	245		250		255		
	TTT GAA ACA ACT TCT GTT ATG ATG CAA GTC AGT TAT TTC AAA CTG ACA							816
	Phe Glu Thr Thr Ser Val Met Met Gln Val Ser Tyr Phe Lys Leu Thr	260		265		270		
30	GGT GGT AAA CGT ATT TTC CGT ATG ACG CCT GTA CAT CAC CAT TTT GAG							864
	Gly Gly Lys Arg Ile Phe Arg Met Thr Pro Val His His His Phe Glu	275		280		285		
35	CTT GGG GGA TTG TCT GGT AAA GGA AAT CCT TGG AGC GAG TGG AAG GTT							912
	Leu Gly Gly Leu Ser Gly Lys Gly Asn Pro Trp Ser Glu Trp Lys Val	290		295		300		
40	GAC TTC TTC TTT TGG GGA GTT GGG CTT CTA GCA AGT CTC CTG ACC CTC							960
	Asp Phe Phe Phe Trp Gly Val Gly Leu Leu Ala Ser Leu Leu Thr Leu	305		310		315		320
45	GCA ATT TTG TAT TTG ATG TAA							981
	Ala Ile Leu Tyr Leu Met	325						

## (2) INFORMATION FOR SEQ ID NO:100:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 326 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

60	Met Phe Ile Ser Ile Ser Ala Gly Ile Val Thr Phe Leu Leu Thr Leu
	1 5 10 15

-176-

Val Gly Ile Pro Ala Phe Ile Gln Phe Tyr Arg Lys Ala Gln Ile Thr  
                     20                    25                    30  
 5 Gly Gln Gln Met His Glu Asp Val Lys Gln His Gln Ala Lys Ala Gly  
                     35                    40                    45  
 Thr Pro Thr Met Gly Gly Leu Val Phe Leu Ile Thr Ser Val Leu Val  
                     50                    55                    60  
 10 Ala Phe Phe Phe Ala Leu Phe Ser Ser Gln Phe Ser Asn Asn Val Gly  
                     65                    70                    75                    80  
 Met Ile Leu Phe Ile Leu Val Leu Tyr Gly Leu Val Gly Phe Leu Asp  
 15                    85                    90                    95  
 Asp Phe Leu Lys Val Phe Arg Lys Ile Asn Glu Gly Leu Asn Pro Lys  
                     100                    105                    110  
 20 Gln Lys Leu Ala Leu Gln Leu Leu Gly Gly Val Ile Phe Tyr Leu Phe  
                     115                    120                    125  
 Tyr Glu Arg Gly Gly Asp Ile Leu Ser Val Phe Gly Tyr Pro Val His  
                     130                    135                    140  
 25 Leu Gly Phe Phe Tyr Ile Phe Phe Ala Leu Phe Trp Leu Val Gly Phe  
                     145                    150                    155                    160  
 Ser Asn Ala Val Asn Leu Thr Asp Gly Val Asp Gly Leu Ala Ser Ile  
 30                    165                    170                    175  
 Ser Val Val Ile Ser Leu Phe Ala Tyr Gly Val Ile Ala Tyr Val Gln  
                     180                    185                    190  
 35 Gly Gln Met Asp Ile Leu Leu Val Ile Leu Ala Met Ile Gly Gly Leu  
                     195                    200                    205  
 Leu Gly Phe Phe Ile Phe Asn His Lys Pro Ala Lys Val Phe Met Gly  
                     210                    215                    220  
 40 Asp Val Gly Ser Leu Ala Leu Gly Gly Met Leu Ala Ala Ile Ser Met  
                     225                    230                    235                    240  
 Ala Leu His Gln Glu Trp Thr Leu Leu Ile Ile Gly Ile Val Tyr Val  
 45                    245                    250                    255  
 Phe Glu Thr Thr Ser Val Met Met Gln Val Ser Tyr Phe Lys Leu Thr  
                     260                    265                    270  
 50 Gly Gly Lys Arg Ile Phe Arg Met Thr Pro Val His His His Phe Glu  
                     275                    280                    285  
 Leu Gly Gly Leu Ser Gly Lys Gly Asn Pro Trp Ser Glu Trp Lys Val  
                     290                    295                    300  
 55 Asp Phe Phe Phe Trp Gly Val Gly Leu Leu Ala Ser Leu Leu Thr Leu  
                     305                    310                    315                    320  
 60 Ala Ile Leu Tyr Leu Met  
                     325

-177-

## (2) INFORMATION FOR SEQ ID NO:101:

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 369 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..366  
 20 (D) OTHER INFORMATION: Dpj

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

25	ATG AGA ATG ATA GTT GGA CAC GGA ATT GAC ATC GAA GAA TTG GCT TCG	48
	Met Arg Met Ile Val Gly His Gly Ile Asp Ile Glu Glu Leu Ala Ser	
	1 5 10 15	
30	ATA GAA AGC GCA GTT ACA CGA CAT GAA GGA TTT GCT AAG CGT GTA CTG	96
	Ile Glu Ser Ala Val Thr Arg His Glu Gly Phe Ala Lys Arg Val Leu	
	20 25 30	
35	ACC GCT CAG GAA ATG GAG CGC TTC ACC AGT CTC AAA GGA CGC AGG CAA	144
	Thr Ala Gln Glu Met Glu Arg Phe Thr Ser Leu Lys Gly Arg Arg Gln	
	35 40 45	
40	ATA GAA TAT TTA GCT GGT CGC TGG TCG GCT AAG GAG GCC TTT TCC AAG	192
	Ile Glu Tyr Leu Ala Gly Arg Trp Ser Ala Lys Glu Ala Phe Ser Lys	
	50 55 60	
45	GCT ATG GGA ACG GGC ATT AGC AAG CTC GGT TTT CAG GAT TTG GAA GTC	240
	Ala Met Gly Thr Gly Ile Ser Lys Leu Gly Phe Gln Asp Leu Glu Val	
	65 70 75 80	
50	TTG AAC AAT GAA CGT GGG GCG CCT TAT TTT AGT CAG GCA CCA TTT TCA	288
	Leu Asn Asn Glu Arg Gly Ala Pro Tyr Phe Ser Gln Ala Pro Phe Ser	
	85 90 95	
55	GGA AAG ATT TGG CTG TCT ATC AGC CAC ACC GAT CAG TTT GTG ACA GCC	336
	Gly Lys Ile Trp Leu Ser Ile Ser His Thr Asp Gln Phe Val Thr Ala	
	100 105 110	
60	AGT GTC ATT TTG GAG GAA AAT CAT GAA AGC TAG	369
	Ser Val Ile Leu Glu Glu Asn His Glu Ser	
	115 120	

## (2) INFORMATION FOR SEQ ID NO:102:

60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 122 amino acids

-178-

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

Met Arg Met Ile Val Gly His Gly Ile Asp Ile Glu Glu Leu Ala Ser  
1 5 10 15  
Ile Glu Ser Ala Val Thr Arg His Glu Gly Phe Ala Lys Arg Val Leu  
20 25 30  
Thr Ala Gln Glu Met Glu Arg Phe Thr Ser Leu Lys Gly Arg Arg Gln  
35 40 45  
Ile Glu Tyr Leu Ala Gly Arg Trp Ser Ala Lys Glu Ala Phe Ser Lys  
50 55 60  
Ala Met Gly Thr Gly Ile Ser Lys Leu Gly Phe Gln Asp Leu Glu Val  
65 70 75 80  
Leu Asn Asn Glu Arg Gly Ala Pro Tyr Phe Ser Gln Ala Pro Phe Ser  
85 90 95  
Gly Lys Ile Trp Leu Ser Ile Ser His Thr Asp Gln Phe Val Thr Ala  
100 105 110  
Ser Val Ile Leu Glu Glu Asn His Glu Ser  
115 120

(2) INFORMATION FOR SEQ ID NO:103:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1260 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45

(iv) ANTI-SENSE: NO

(ix) FEATURE:

50

(A) NAME/KEY: CDS  
(B) LOCATION: 1..1260  
(D) OTHER INFORMATION: MurZ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

55

ATG AGA AAA ATT GTT ATC AAT GGT GGA TTA CCA CTG CAA GGT GAA ATC 48  
Met Arg Lys Ile Val Ile Asn Gly Gly Leu Pro Leu Gln Gly Glu Ile  
1 5 10 15

60

ACT ATT AGT GGT GCT AAA AAT AGT GTC GTT GCC TTA ATT CCA GCT ATT 96  
Thr Ile Ser Gly Ala Lys Asn Ser Val Val Ala Leu Ile Pro Ala Ile

-179-

	20	25	30	
5	ATC TTG GCT GAT GAT GTG GTG ACT TTG GAT TGC GTT CCA GAT ATT TCG Ile Leu Ala Asp Asp Val Val Thr Leu Asp Cys Val Pro Asp Ile Ser 35 40 45	144		
10	GAT GTA GCC AGT CTT GTC GAA ATC ATG GAA TTG ATG GGA GCT ACT GTT Asp Val Ala Ser Leu Val Glu Ile Met Glu Leu Met Gly Ala Thr Val 50 55 60	192		
15	AAG CGT TAT GAC GAT GTA TTG GAG ATT GAC CCA AGA GGT GTT CAA AAT Lys Arg Tyr Asp Asp Val Leu Glu Ile Asp Pro Arg Gly Val Gln Asn 65 70 75 80	240		
20	ATT CCA ATG CCT TAT GGT AAA ATT AAC AGT CTT CGT GCA TCT TAC TAT Ile Pro Met Pro Tyr Gly Lys Ile Asn Ser Leu Arg Ala Ser Tyr Tyr 85 90 95	288		
25	TTT TAT GGG AGC CTC TTA GGC CGT TTT GGT GAA GCG ACA GTT GGT CTA Phe Tyr Gly Ser Leu Leu Gly Arg Phe Gly Glu Ala Thr Val Gly Leu 100 105 110	336		
30	CCG GGA GGA TGT GAT CTT GGT CCT CGT CCG ATT GAC TTA CAC CTT AAG Pro Gly Gly Cys Asp Leu Gly Pro Arg Pro Ile Asp Leu His Leu Lys 115 120 125	384		
35	GCG TTT GAA GCT ATG GGT GCC ACT GCT AGC TAC GAG GGA GAT AAC ATG Ala Phe Glu Ala Met Gly Ala Thr Ala Ser Tyr Glu Gly Asp Asn Met 130 135 140	432		
40	AAG TTA TCT GCT AAA GAT ACA GGA CTT CAT GGT GCA AGT ATT TAC ATG Lys Leu Ser Ala Lys Asp Thr Gly Leu His Gly Ala Ser Ile Tyr Met 145 150 155 160	480		
45	GAT ACG GTT AGT GTG GGA GCA ACG ATT AAT ACG ATG ATT GCT GCG GTT Asp Thr Val Ser Val Gly Ala Thr Ile Asn Thr Met Ile Ala Ala Val 165 170 175	528		
50	AAA GCA AAT GGT CGT ACT ATT ATT GAA AAT GCA GCC CGT GAA CCT GAG Lys Ala Asn Gly Arg Thr Ile Ile Glu Asn Ala Ala Arg Glu Pro Glu 180 185 190	576		
55	ATT ATT GAT GTA GCT ACT CTC TTG AAT AAT ATG GGT GCC CAT ATC CGT Ile Ile Asp Val Ala Thr Leu Leu Asn Asn Met Gly Ala His Ile Arg 195 200 205	624		
60	GGG GCA GGA ACT AAT ATC ATC ATT ATT GAT GGT GTT GAA AGA TTA CAT Gly Ala Gly Thr Asn Ile Ile Ile Ile Asp Gly Val Glu Arg Leu His 210 215 220	672		
65	GGG ACA CGT CAT CAG GTG ATT CCA GAC CGC ATT GAA GCT GGA ACA TAT Gly Thr Arg His Gln Val Ile Pro Asp Arg Ile Glu Ala Gly Thr Tyr 225 230 235 240	720		
70	ATA TCT TTA GCT GCT GCA GTT GGT AAA GGA ATT CGT ATA AAT AAT GTT Ile Ser Leu Ala Ala Ala Val Gly Lys Gly Ile Arg Ile Asn Asn Val 245 250 255	768		
75	CTT TAC GAA CAC CTG GAA GGG TTT GTT GCT AAG TTG GAA GAA ATG GGA Leu Tyr Glu His Leu Glu Gly Phe Val Ala Lys Leu Glu Glu Met Gly 260 265 270	816		



-180-

5 GTG AGA ATG ACT GTA TCT GAA GAC AGC ATT TTT GTC GAG GAA CAG TCT 864  
 Val Arg Met Thr Val Ser Glu Asp Ser Ile Phe Val Glu Glu Gln Ser  
 275 280 285  
 AAT TTG AAA GCA ATC AAT ATT AAG ACA GCT CCT TAC CCA GGC TTT GCA 912  
 Asn Leu Lys Ala Ile Asn Ile Lys Thr Ala Pro Tyr Pro Gly Phe Ala  
 290 295 300  
 10 ACT GAT TTG CAA CAA CCG CTT ACC CCT CTT TTA CTA AGA GCG AAT GGT 960  
 Thr Asp Leu Gln Gln Pro Leu Thr Pro Leu Leu Leu Arg Ala Asn Gly  
 305 310 315 320  
 15 CGT GGT ACA ATT GTC GAT ACG ATT TAC GAA AAA CGT GTA AAT CAT GTT 1008  
 Arg Gly Thr Ile Val Asp Thr Ile Tyr Glu Lys Arg Val Asn His Val  
 325 330 335  
 20 TTT GAA CTA GCA AAG ATG GAT GCG GAT ATT TCG ACA ACA AAT GGT CAT 1056  
 Phe Glu Leu Ala Lys Met Asp Ala Asp Ile Ser Thr Thr Asn Gly His  
 340 345 350  
 25 ATT TTG TAC ACG GGT GGA CGT GAT TTA CGT GGG GCC AGT GTT AAA GCG 1104  
 Ile Leu Tyr Thr Gly Gly Arg Asp Leu Arg Gly Ala Ser Val Lys Ala  
 355 360 365  
 ACC GAC TTA AGA GCT GGG GCT GCA CTA GTC ATT GCT GGG CTT ATG GCT 1152  
 Thr Asp Leu Arg Ala Gly Ala Ala Leu Val Ile Ala Gly Leu Met Ala  
 370 375 380  
 30 GAA GGC AAA ACT GAA ATT ACC AAT ATC GAG TTT ATC TTA CGT GGT TAT 1200  
 Glu Gly Lys Thr Glu Ile Thr Asn Ile Glu Phe Ile Leu Arg Gly Tyr  
 385 390 395 400  
 35 TCT GAT ATT ATC GAA AAA TTA CGT AAT TTA GGA GCG GAT ATT AGA CTT 1248  
 Ser Asp Ile Ile Glu Lys Leu Arg Asn Leu Gly Ala Asp Ile Arg Leu  
 405 410 415  
 40 GTT GAG GAT TAA 1260  
 Val Glu Asp  
 419

## (2) INFORMATION FOR SEQ ID NO:104:

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 419 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

55 Met Arg Lys Ile Val Ile Asn Gly Gly Leu Pro Leu Gln Gly Glu Ile  
 1 5 10 15  
 Thr Ile Ser Gly Ala Lys Asn Ser Val Val Ala Leu Ile Pro Ala Ile  
 20 25 30  
 60 Ile Leu Ala Asp Asp Val Val Thr Leu Asp Cys Val Pro Asp Ile Ser  
 35 40 45

-181-

Asp Val Ala Ser Leu Val Glu Ile Met Glu Leu Met Gly Ala Thr Val  
 50 55 60  
 5 Lys Arg Tyr Asp Asp Val Leu Glu Ile Asp Pro Arg Gly Val Gln Asn  
 65 70 75 80  
 Ile Pro Met Pro Tyr Gly Lys Ile Asn Ser Leu Arg Ala Ser Tyr Tyr  
 85 90 95  
 10 Phe Tyr Gly Ser Leu Leu Gly Arg Phe Gly Glu Ala Thr Val Gly Leu  
 100 105 110  
 15 Pro Gly Gly Cys Asp Leu Gly Pro Arg Pro Ile Asp Leu His Leu Lys  
 115 120 125  
 Ala Phe Glu Ala Met Gly Ala Thr Ala Ser Tyr Glu Gly Asp Asn Met  
 130 135 140  
 20 Lys Leu Ser Ala Lys Asp Thr Gly Leu His Gly Ala Ser Ile Tyr Met  
 145 150 155 160  
 Asp Thr Val Ser Val Gly Ala Thr Ile Asn Thr Met Ile Ala Ala Val  
 165 170 175  
 25 Lys Ala Asn Gly Arg Thr Ile Ile Glu Asn Ala Ala Arg Glu Pro Glu  
 180 185 190  
 30 Ile Ile Asp Val Ala Thr Leu Leu Asn Asn Met Gly Ala His Ile Arg  
 195 200 205  
 Gly Ala Gly Thr Asn Ile Ile Ile Ile Asp Gly Val Glu Arg Leu His  
 210 215 220  
 35 Gly Thr Arg His Gln Val Ile Pro Asp Arg Ile Glu Ala Gly Thr Tyr  
 225 230 235 240  
 Ile Ser Leu Ala Ala Ala Val Gly Lys Gly Ile Arg Ile Asn Asn Val  
 245 250 255  
 40 Leu Tyr Glu His Leu Glu Gly Phe Val Ala Lys Leu Glu Glu Met Gly  
 260 265 270  
 45 Val Arg Met Thr Val Ser Glu Asp Ser Ile Phe Val Glu Glu Gln Ser  
 275 280 285  
 Asn Leu Lys Ala Ile Asn Ile Lys Thr Ala Pro Tyr Pro Gly Phe Ala  
 290 295 300  
 50 Thr Asp Leu Gln Gln Pro Leu Thr Pro Leu Leu Leu Arg Ala Asn Gly  
 305 310 315 320  
 Arg Gly Thr Ile Val Asp Thr Ile Tyr Glu Lys Arg Val Asn His Val  
 325 330 335  
 55 Phe Glu Leu Ala Lys Met Asp Ala Asp Ile Ser Thr Thr Asn Gly His  
 340 345 350  
 60 Ile Leu Tyr Thr Gly Gly Arg Asp Leu Arg Gly Ala Ser Val Lys Ala  
 355 360 365

-182-

Thr Asp Leu Arg Ala Gly Ala Ala Leu Val Ile Ala Gly Leu Met Ala  
 370 375 380

5 Glu Gly Lys Thr Glu Ile Thr Asn Ile Glu Phe Ile Leu Arg Gly Tyr  
 385 390 395 400

Ser Asp Ile Ile Glu Lys Leu Arg Asn Leu Gly Ala Asp Ile Arg Leu  
 405 410 415

10 Val Glu Asp  
 419

## (2) INFORMATION FOR SEQ ID NO:105:

15

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1008 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

20

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25

(iv) ANTI-SENSE: NO

## (ix) FEATURE:

30

(A) NAME/KEY: CDS

(B) LOCATION: 1..1008

(D) OTHER INFORMATION: FtsZ

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

	ATG ACA TTT TCA TTT GAT ACA GCT GCT GCT CAA GGG GCA GTG ATT AAA	48
	Met Thr Phe Ser Phe Asp Thr Ala Ala Ala Gln Gly Ala Val Ile Lys	
40	1 5 10 15	
	GTA ATT GGT GTC GGT GGA GGT GGT GGC AAT GCC ATC AAC CGT ATG GTC	96
	Val Ile Gly Val Gly Gly Gly Gly Gly Asn Ala Ile Asn Arg Met Val	
	20 25 30	
45	GAC GAA GGT GTT ACA GGC GTA GAA TTT ATC GCA GCA AAC ACA GAT GTA	144
	Asp Glu Gly Val Thr Gly Val Glu Phe Ile Ala Ala Asn Thr Asp Val	
	35 40 45	
50	CAA GCA TTG AGT AGT ACA AAA GCT GAG ACT GTT ATT CAG TTG GGA CCT	192
	Gln Ala Leu Ser Ser Thr Lys Ala Glu Thr Val Ile Gln Leu Gly Pro	
	50 55 60	
55	AAA TTG ACT CGT GGT TTG GGT GCA GGA GGT CAA CCT GAG GTT GGT CGT	240
	Lys Leu Thr Arg Gly Leu Gly Ala Gly Gly Gln Pro Glu Val Gly Arg	
	65 70 75 80	
	AAA GCC GCT GAA GAA AGC GAA GAA ACA CTG ACG GAA GCT ATT AGT GGT	288
	Lys Ala Ala Glu Glu Ser Glu Glu Thr Leu Thr Glu Ala Ile Ser Gly	
	85 90 95	
60	GCC GAT ATG GTC TTC ATC ACT GCT GGT ATG GGA GGA GGC TCT GGA ACT	336

-183-

	Ala	Asp	Met	Val	Phe	Ile	Thr	Ala	Gly	Met	Gly	Gly	Gly	Ser	Gly	Thr	
				100					105					110			
5	GGA	GCT	GCT	CCT	GTT	ATT	GCT	CGT	ATC	GCC	AAA	GAT	TTA	GGT	GCG	CTT	384
	Gly	Ala	Ala	Pro	Val	Ile	Ala	Arg	Ile	Ala	Lys	Asp	Leu	Gly	Ala	Leu	
			115					120					125				
10	ACA	GTT	GGT	GTT	GTA	ACA	CGT	CCC	TTT	GGT	TTT	GAA	GGA	AGT	AAG	CGT	432
	Thr	Val	Gly	Val	Val	Thr	Arg	Pro	Phe	Gly	Phe	Glu	Gly	Ser	Lys	Arg	
		130					135					140					
15	GGA	CAA	TTT	GCT	GTA	GAA	GGA	ATC	AAT	CAA	CTT	CGT	GAG	CAT	GTA	GAC	480
	Gly	Gln	Phe	Ala	Val	Glu	Gly	Ile	Asn	Gln	Leu	Arg	Glu	His	Val	Asp	
	145					150					155					160	
20	ACT	CTA	TTG	ATT	ATC	TCA	AAC	AAC	AAT	TTG	CTT	GAA	ATT	GTT	GAT	AAG	528
	Thr	Leu	Leu	Ile	Ile	Ser	Asn	Asn	Asn	Leu	Leu	Glu	Ile	Val	Asp	Lys	
					165					170					175		
25	AAA	ACA	CCG	CTT	TTG	GAG	GCT	CTT	AGC	GAA	GCG	GAT	AAC	GTT	CTT	CGT	576
	Lys	Thr	Pro	Leu	Leu	Glu	Ala	Leu	Ser	Glu	Ala	Asp	Asn	Val	Leu	Arg	
				180					185					190			
30	CAA	GGT	GTT	CAA	GGG	ATT	ACC	GAT	TTG	ATT	ACC	AAT	CCA	GGA	TTG	ATT	624
	Gln	Gly	Val	Gln	Gly	Ile	Thr	Asp	Leu	Ile	Thr	Asn	Pro	Gly	Leu	Ile	
			195					200					205				
35	AAC	CTT	GAC	TTT	GCC	GAT	GTG	AAA	ACG	GTA	ATG	GCA	AAC	AAA	GGG	AAT	672
	Asn	Leu	Asp	Phe	Ala	Asp	Val	Lys	Thr	Val	Met	Ala	Asn	Lys	Gly	Asn	
		210					215					220					
40	GCT	CTT	ATG	GGT	ATT	GGT	ATC	GGT	AGT	GGA	GAA	GAA	CGT	GTG	GTA	GAA	720
	Ala	Leu	Met	Gly	Ile	Gly	Ile	Gly	Ser	Gly	Glu	Glu	Arg	Val	Val	Glu	
	225					230					235					240	
45	GCG	GCA	CGT	AAG	GCA	ATC	TAT	TCA	CCA	CTT	CTT	GAA	ACA	ACT	ATT	GAC	768
	Ala	Ala	Arg	Lys	Ala	Ile	Tyr	Ser	Pro	Leu	Leu	Glu	Thr	Thr	Ile	Asp	
					245					250					255		
50	GGT	GCT	GAG	GAT	GTT	ATC	GTC	AAC	GTT	ACT	GGT	GGT	CTT	GAC	TTA	ACC	816
	Gly	Ala	Glu	Asp	Val	Ile	Val	Asn	Val	Thr	Gly	Gly	Leu	Asp	Leu	Thr	
				260					265					270			
55	TTG	ATT	GAG	GCA	GAA	GAG	GCT	TCA	CAA	ATT	GTG	AAC	CAG	GCA	GCA	GGT	864
	Leu	Ile	Glu	Ala	Glu	Glu	Ala	Ser	Gln	Ile	Val	Asn	Gln	Ala	Ala	Gly	
			275					280					285				
60	CAA	GGA	GTG	AAC	ATC	TGG	CTC	GGT	ACT	TCA	ATT	GAT	GAA	AGT	ATG	CGT	912
	Gln	Gly	Val	Asn	Ile	Trp	Leu	Gly	Thr	Ser	Ile	Asp	Glu	Ser	Met	Arg	
		290					295					300					
65	GAT	GAA	ATT	CGT	GTA	ACA	GTT	GTC	GCA	ACG	GGT	GTT	CGT	CAA	GAC	CGC	960
	Asp	Glu	Ile	Arg	Val	Thr	Val	Val	Ala	Thr	Gly	Val	Arg	Gln	Asp	Arg	
	305					310					315					320	
70	GTA	GAA	AAG	GTT	GTG	GCT	CCA	CAA	GCT	AGA	TCA	CCG	CGC	CTA	GGA	TAA	1008
	Val	Glu	Lys	Val	Val	Ala	Pro	Gln	Ala	Arg	Ser	Pro	Arg	Leu	Gly	*	
					325					330					335		

(2) INFORMATION FOR SEQ ID NO:106:

-184-

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 336 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

5 Met Thr Phe Ser Phe Asp Thr Ala Ala Ala Gln Gly Ala Val Ile Lys  
 1 5 10 15  
 10 Val Ile Gly Val Gly Gly Gly Gly Gly Asn Ala Ile Asn Arg Met Val  
 15 20 25 30  
 Asp Glu Gly Val Thr Gly Val Glu Phe Ile Ala Ala Asn Thr Asp Val  
 35 40 45  
 20 Gln Ala Leu Ser Ser Thr Lys Ala Glu Thr Val Ile Gln Leu Gly Pro  
 50 55 60  
 Lys Leu Thr Arg Gly Leu Gly Ala Gly Gly Gln Pro Glu Val Gly Arg  
 65 70 75 80  
 25 Lys Ala Ala Glu Glu Ser Glu Glu Thr Leu Thr Glu Ala Ile Ser Gly  
 85 90 95  
 30 Ala Asp Met Val Phe Ile Thr Ala Gly Met Gly Gly Gly Ser Gly Thr  
 100 105 110  
 Gly Ala Ala Pro Val Ile Ala Arg Ile Ala Lys Asp Leu Gly Ala Leu  
 115 120 125  
 35 Thr Val Gly Val Val Thr Arg Pro Phe Gly Phe Glu Gly Ser Lys Arg  
 130 135 140  
 Gly Gln Phe Ala Val Glu Gly Ile Asn Gln Leu Arg Glu His Val Asp  
 145 150 155 160  
 40 Thr Leu Leu Ile Ile Ser Asn Asn Asn Leu Leu Glu Ile Val Asp Lys  
 165 170 175  
 Lys Thr Pro Leu Leu Glu Ala Leu Ser Glu Ala Asp Asn Val Leu Arg  
 180 185 190  
 45 Gln Gly Val Gln Gly Ile Thr Asp Leu Ile Thr Asn Pro Gly Leu Ile  
 195 200 205  
 50 Asn Leu Asp Phe Ala Asp Val Lys Thr Val Met Ala Asn Lys Gly Asn  
 210 215 220  
 Ala Leu Met Gly Ile Gly Ile Gly Ser Gly Glu Glu Arg Val Val Glu  
 225 230 235 240  
 55 Ala Ala Arg Lys Ala Ile Tyr Ser Pro Leu Leu Glu Thr Thr Ile Asp  
 245 250 255  
 60 Gly Ala Glu Asp Val Ile Val Asn Val Thr Gly Gly Leu Asp Leu Thr  
 260 265 270

-185-

Leu Ile Glu Ala Glu Glu Ala Ser Gln Ile Val Asn Gln Ala Ala Gly  
                   275                  280                  285

5 Gln Gly Val Asn Ile Trp Leu Gly Thr Ser Ile Asp Glu Ser Met Arg  
                   290                  295                  300

Asp Glu Ile Arg Val Thr Val Val Ala Thr Gly Val Arg Gln Asp Arg  
   305                  310                  315                  320

10 Val Glu Lys Val Val Ala Pro Gln Ala Arg Ser Pro Arg Leu Gly  
                   325                  330                  335

## (2) INFORMATION FOR SEQ ID NO:107:

15 (i) SEQUENCE CHARACTERISTICS:  
       (A) LENGTH: 525 base pairs  
       (B) TYPE: nucleic acid  
       (C) STRANDEDNESS: single  
       (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(ix) FEATURE:  
       (A) NAME/KEY: CDS  
       (B) LOCATION: 1..525  
       (D) OTHER INFORMATION: grpE

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

35	ATG GCC CAA GAT ATA AAA AAT GAA GAA GTA GAA GAA GTT CAA GAA GAG	48
	Met Ala Gln Asp Ile Lys Asn Glu Glu Val Glu Glu Val Gln Glu Glu	
	1                  5                  10                  15	
40	GAA GTT GTG GAA ACA GCT GAA GAA ACA ACT CCT GAA AAG TCT GAG TTG	96
	Glu Val Val Glu Thr Ala Glu Glu Thr Thr Pro Glu Lys Ser Glu Leu	
	20                  25                  30	
45	GAC TTG GCA AAT GAA CGT GCA GAT GAG TTC GAA AAC AAA TAT CTT CGC	144
	Asp Leu Ala Asn Glu Arg Ala Asp Glu Phe Glu Asn Lys Tyr Leu Arg	
	35                  40                  45	
50	GCT CAT GCA GAA ATG CAA AAT ATC CAA CGC CGT GCC AAT GAA GAA CGT	192
	Ala His Ala Glu Met Gln Asn Ile Gln Arg Arg Ala Asn Glu Glu Arg	
	50                  55                  60	
	CAA AAC TTG CAA CGT TAT CGT AGC CAG GAC TTG GCA AAA GCA ATC TTA	240
	Gln Asn Leu Gln Arg Tyr Arg Ser Gln Asp Leu Ala Lys Ala Ile Leu	
	65                  70                  75                  80	
55	CCA TCT CTT GAC AAC CTT GAG CGT GCA CTT GCA GTT GAA GGT TTG ACA	288
	Pro Ser Leu Asp Asn Leu Glu Arg Ala Leu Ala Val Glu Gly Leu Thr	
	85                  90                  95	
60	GAT GAT GTG AAG AAG GGC TTG GCG ATG GTG CAA GAA AGC TTG ATT CAC	336
	Asp Asp Val Lys Lys Gly Leu Ala Met Val Gln Glu Ser Leu Ile His	

-186-

	100	105	110	
5	GCT TTG AAA GAA GAA GGA ATT GAA GAA ATC GCA GCA GAT GGC GAA TTT Ala Leu Lys Glu Glu Gly Ile Glu Glu Ile Ala Ala Asp Gly Glu Phe 115 120 125			384
10	GAC CAT AAC TAC CAT ATG GCC ATC CAA ACT CTC CCA GGA GAC GAT GAA Asp His Asn Tyr His Met Ala Ile Gln Thr Leu Pro Gly Asp Asp Glu 130 135 140			432
15	CAC CCA GTA GAT ACC ATC GCC CAA GTC TTT CAA AAA GGC TAC AAA CTC His Pro Val Asp Thr Ile Ala Gln Val Phe Gln Lys Gly Tyr Lys Leu 145 150 155 160			480
20	CAT GAC CGC ATC CTA CGC CCA GCA ATG GTA GTG GTG TAT AAC TAA His Asp Arg Ile Leu Arg Pro Ala Met Val Val Val Tyr Asn * 165 170 174			525
(2) INFORMATION FOR SEQ ID NO:108:				
(i) SEQUENCE CHARACTERISTICS:				
(A) LENGTH: 174 amino acids				
(B) TYPE: amino acid				
(D) TOPOLOGY: linear				
(ii) MOLECULE TYPE: protein				
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:				
30	Met Ala Gln Asp Ile Lys Asn Glu Glu Val Glu Glu Val Gln Glu Glu 1 5 10 15			
35	Glu Val Val Glu Thr Ala Glu Glu Thr Thr Pro Glu Lys Ser Glu Leu 20 25 30			
40	Asp Leu Ala Asn Glu Arg Ala Asp Glu Phe Glu Asn Lys Tyr Leu Arg 35 40 45			
45	Ala His Ala Glu Met Gln Asn Ile Gln Arg Arg Ala Asn Glu Glu Arg 50 55 60			
50	Gln Asn Leu Gln Arg Tyr Arg Ser Gln Asp Leu Ala Lys Ala Ile Leu 65 70 75 80			
55	Pro Ser Leu Asp Asn Leu Glu Arg Ala Leu Ala Val Glu Gly Leu Thr 85 90 95			
60	Asp Asp Val Lys Lys Gly Leu Ala Met Val Gln Glu Ser Leu Ile His 100 105 110			
	Ala Leu Lys Glu Glu Gly Ile Glu Glu Ile Ala Ala Asp Gly Glu Phe 115 120 125			
	Asp His Asn Tyr His Met Ala Ile Gln Thr Leu Pro Gly Asp Asp Glu 130 135 140			
	His Pro Val Asp Thr Ile Ala Gln Val Phe Gln Lys Gly Tyr Lys Leu 145 150 155 160			
	His Asp Arg Ile Leu Arg Pro Ala Met Val Val Val Tyr Asn			

-187-

165

170

174

## (2) INFORMATION FOR SEQ ID NO:109:

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 582 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..582  
 (D) OTHER INFORMATION: HI1648

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

25	ATG AAA ATC GGA ATA TTG GCC TTG CAA GGG GCC TTT GCA GAA CAT GCA	48
	Met Lys Ile Gly Ile Leu Ala Leu Gln Gly Ala Phe Ala Glu His Ala	
	1 5 10 15	
30	AAA GTG CTA GAT CAA TTA GGT GTC GAG AGT GTA GAA CTC AGA AAT CTA	96
	Lys Val Leu Asp Gln Leu Gly Val Glu Ser Val Glu Leu Arg Asn Leu	
	20 25 30	
35	GAT GAT TTT CAG CAA GAT CAG AGT GAC TTG TCG GGT TTG ATT TTG CCT	144
	Asp Asp Phe Gln Gln Asp Gln Ser Asp Leu Ser Gly Leu Ile Leu Pro	
	35 40 45	
40	GGT GGT GAG TCT ACA ACC ATG GGC AAG CTC TTA CGT GAC CAG AAC ATG	192
	Gly Gly Glu Ser Thr Thr Met Gly Lys Leu Leu Arg Asp Gln Asn Met	
	50 55 60	
45	CTA CTT CCC ATA CGA GAA GCC ATT CTA TCT GGC TTA CCA GTG TTT GGG	240
	Leu Leu Pro Ile Arg Glu Ala Ile Leu Ser Gly Leu Pro Val Phe Gly	
	65 70 75 80	
	ACC TGT GCG GGC TTA ATT TTG CTG GCT AAG GAA ATC ACT TCT CAG AAA	288
	Thr Cys Ala Gly Leu Ile Leu Leu Ala Lys Glu Ile Thr Ser Gln Lys	
	85 90 95	
50	GAG AGT CAT CTA GGA ACT ATG GAT ATG GTG GTC GAG CGT AAT GCT TAT	336
	Glu Ser His Leu Gly Thr Met Asp Met Val Val Glu Arg Asn Ala Tyr	
	100 105 110	
55	GGG CGC CAA TTA GGA AGT TTC TAC ACG GAA GCA GAA TGT AAG GGA GTT	384
	Gly Arg Gln Leu Gly Ser Phe Tyr Thr Glu Ala Glu Cys Lys Gly Val	
	115 120 125	
60	GGC AAG ATT CCA ATG ACC TTT ATC CGT GGT CCG ATT ATC AGT AGT GTT	432
	Gly Lys Ile Pro Met Thr Phe Ile Arg Gly Pro Ile Ile Ser Ser Val	
	130 135 140	



-188-

GGT GAG GGT GTA GAA ATT TTA GCA ATA GTG AAC AAT CAA ATT GTT GCA 480  
 Gly Glu Gly Val Glu Ile Leu Ala Ile Val Asn Asn Gln Ile Val Ala  
 145 150 155 160

5 GCC CAA GAA AAA AAT ATG TTG GTA AGT TCT TTT CAT CCA GAA TTG ACT 528  
 Ala Gln Glu Lys Asn Met Leu Val Ser Ser Phe His Pro Glu Leu Thr  
 165 170 175

10 GAT GAT GTG CGC TTG CAC CAG TAC TTT ATC AAT ATG TGT AAA GAA AAA 576  
 Asp Asp Val Arg Leu His Gln Tyr Phe Ile Asn Met Cys Lys Glu Lys  
 180 185 190

AGT TGA 582  
 Ser \*

15

## (2) INFORMATION FOR SEQ ID NO:110:

- 20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 194 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear
- 25 (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

30 Met Lys Ile Gly Ile Leu Ala Leu Gln Gly Ala Phe Ala Glu His Ala  
 1 5 10 15

Lys Val Leu Asp Gln Leu Gly Val Glu Ser Val Glu Leu Arg Asn Leu  
 20 25 30

35 Asp Asp Phe Gln Gln Asp Gln Ser Asp Leu Ser Gly Leu Ile Leu Pro  
 35 40 45

Gly Gly Glu Ser Thr Thr Met Gly Lys Leu Leu Arg Asp Gln Asn Met  
 50 55 60

40 Leu Leu Pro Ile Arg Glu Ala Ile Leu Ser Gly Leu Pro Val Phe Gly  
 65 70 75 80

45 Thr Cys Ala Gly Leu Ile Leu Leu Ala Lys Glu Ile Thr Ser Gln Lys  
 85 90 95

Glu Ser His Leu Gly Thr Met Asp Met Val Val Glu Arg Asn Ala Tyr  
 100 105 110

50 Gly Arg Gln Leu Gly Ser Phe Tyr Thr Glu Ala Glu Cys Lys Gly Val  
 115 120 125

Gly Lys Ile Pro Met Thr Phe Ile Arg Gly Pro Ile Ile Ser Ser Val  
 130 135 140

55 Gly Glu Gly Val Glu Ile Leu Ala Ile Val Asn Asn Gln Ile Val Ala  
 145 150 155 160

60 Ala Gln Glu Lys Asn Met Leu Val Ser Ser Phe His Pro Glu Leu Thr  
 165 170 175

-189-

Asp Asp Val Arg Leu His Gln Tyr Phe Ile Asn Met Cys Lys Glu Lys  
 180 185 190

Ser

5

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 546 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20 (iv) ANTI-SENSE: NO

(ix) FEATURE:

- 25 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..543  
 (D) OTHER INFORMATION: pgsA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

30	ATG AAA AAA GAA CAA ATT CCC AAT CTC TTA ACA ATA GGT CGA ATT CTC	48
	Met Lys Lys Glu Gln Ile Pro Asn Leu Leu Thr Ile Gly Arg Ile Leu	
	1 5 10 15	
35	TTT ATA CCT ATT TTT ATC TTT ATT TTA ACG ATA GGA AAT TCG ATA GAG	96
	Phe Ile Pro Ile Phe Ile Phe Ile Leu Thr Ile Gly Asn Ser Ile Glu	
	20 25 30	
40	AGT CAT ATA GTT GCA GCT ATT ATC TTT GCT GTT GCC AGT ATT ACC GAC	144
	Ser His Ile Val Ala Ala Ile Ile Phe Ala Val Ala Ser Ile Thr Asp	
	35 40 45	
45	TAT TTA GAT GGA TAT TTA GCT CGT AAA TGG AAT GTG GTC AGT AAT TTT	192
	Tyr Leu Asp Gly Tyr Leu Ala Arg Lys Trp Asn Val Val Ser Asn Phe	
	50 55 60	
50	GGT AAA TTT GCA GAT CCT ATG GCG GAT AAG TTA CTA GTT ATG TCG GCT	240
	Gly Lys Phe Ala Asp Pro Met Ala Asp Lys Leu Leu Val Met Ser Ala	
	65 70 75 80	
55	TTT ATT ATG TTG ATT GAG TTA GGT ATG GCT CCG GCT TGG ATT GTT GCA	288
	Phe Ile Met Leu Ile Glu Leu Gly Met Ala Pro Ala Trp Ile Val Ala	
	85 90 95	
60	GTG ATT ATC TGT CGT GAG TTA GCT GTG ACA GGT TTA AGG CTT TTA TTG	336
	Val Ile Ile Cys Arg Glu Leu Ala Val Thr Gly Leu Arg Leu Leu Leu	
	100 105 110	
60	GTT GAA ACT GGT GGA ACA ATT TTA GCA GCA GCA ATG CCT GGA AAA ATT	384
	Val Glu Thr Gly Gly Thr Ile Leu Ala Ala Ala Met Pro Gly Lys Ile	
	115 120 125	

-190-

	AAA ACT TTT AGT CAG ATG TTT GCT ATT ATT TTC TTG CTA TTA CAT TGG	432
	Lys Thr Phe Ser Gln Met Phe Ala Ile Ile Phe Leu Leu Leu His Trp	
	130 135 140	
5	ACT TTG CTT GGT CAA GTT CTA CTT TAT GTA GCC TTA TTT TTC ACT ATC	480
	Thr Leu Leu Gly Gln Val Leu Leu Tyr Val Ala Leu Phe Phe Thr Ile	
	145 150 155 160	
10	TAC TCT GGC TAT GAC TAT TTC AAG GGT AGT GCC TAT GTA TTT AAA GGG	528
	Tyr Ser Gly Tyr Asp Tyr Phe Lys Gly Ser Ala Tyr Val Phe Lys Gly	
	165 170 175	
	ACA TTT GGT TCG AAA TGA	546
15	Thr Phe Gly Ser Lys	
	180	

## (2) INFORMATION FOR SEQ ID NO:112:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 181 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

30	Met Lys Lys Glu Gln Ile Pro Asn Leu Leu Thr Ile Gly Arg Ile Leu	
	1 5 10 15	
	Phe Ile Pro Ile Phe Ile Phe Ile Leu Thr Ile Gly Asn Ser Ile Glu	
	20 25 30	
35	Ser His Ile Val Ala Ala Ile Ile Phe Ala Val Ala Ser Ile Thr Asp	
	35 40 45	
40	Tyr Leu Asp Gly Tyr Leu Ala Arg Lys Trp Asn Val Val Ser Asn Phe	
	50 55 60	
	Gly Lys Phe Ala Asp Pro Met Ala Asp Lys Leu Leu Val Met Ser Ala	
	65 70 75 80	
45	Phe Ile Met Leu Ile Glu Leu Gly Met Ala Pro Ala Trp Ile Val Ala	
	85 90 95	
	Val Ile Ile Cys Arg Glu Leu Ala Val Thr Gly Leu Arg Leu Leu Leu	
	100 105 110	
50	Val Glu Thr Gly Gly Thr Ile Leu Ala Ala Ala Met Pro Gly Lys Ile	
	115 120 125	
	Lys Thr Phe Ser Gln Met Phe Ala Ile Ile Phe Leu Leu Leu His Trp	
	130 135 140	
55	Thr Leu Leu Gly Gln Val Leu Leu Tyr Val Ala Leu Phe Phe Thr Ile	
	145 150 155 160	
60	Tyr Ser Gly Tyr Asp Tyr Phe Lys Gly Ser Ala Tyr Val Phe Lys Gly	
	165 170 175	

-191-

Thr Phe Gly Ser Lys  
180

## (2) INFORMATION FOR SEQ ID NO:113:

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1224 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..1221  
 (D) OTHER INFORMATION: RodA

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

25	ATG AAA CGT TCT CTC GAC TCT AGA GTC GAT TAT AGT TTG CTC TTG CCA	48
	Met Lys Arg Ser Leu Asp Ser Arg Val Asp Tyr Ser Leu Leu Leu Pro	
	1 5 10 15	
30	GTA TTT TTT CTA CTG GTC ATC GGT GTG GTG GCT ATC TAT ATA GCC GTT	96
	Val Phe Phe Leu Leu Val Ile Gly Val Val Ala Ile Tyr Ile Ala Val	
	20 25 30	
35	AGT CAT GAT TAT CCC AAT AAT ATT CTG CCC ATT TTA GGG CAG CAG GTC	144
	Ser His Asp Tyr Pro Asn Asn Ile Leu Pro Ile Leu Gly Gln Gln Val	
	35 40 45	
40	GCC TGG ATT GCC TTG GGG CTT GTG ATT GGT TTT GTG GTC ATG CTC TTT	192
	Ala Trp Ile Ala Leu Gly Leu Val Ile Gly Phe Val Val Met Leu Phe	
	50 55 60	
45	AAT ACA GAA TTT CTT TGG AAG GTG ACC CCC TTT CTA TAT ATT TTA GGC	240
	Asn Thr Glu Phe Leu Trp Lys Val Thr Pro Phe Leu Tyr Ile Leu Gly	
	65 70 75 80	
50	TTG GGA CTT ATG ATC TTG CCG ATT GTA TTT TAT AAT CCA AGC TTA GTT	288
	Leu Gly Leu Met Ile Leu Pro Ile Val Phe Tyr Asn Pro Ser Leu Val	
	85 90 95	
55	GCA TCA ACG GGT GCC AAA AAC TGG GTA TCA ATA AAT GGA ATT ACC CTA	336
	Ala Ser Thr Gly Ala Lys Asn Trp Val Ser Ile Asn Gly Ile Thr Leu	
	100 105 110	
60	TTT CAA CCG TCA GAA TTT ATG AAG ATA TCC TAT ATC CTC ATG TTG GCT	384
	Phe Gln Pro Ser Glu Phe Met Lys Ile Ser Tyr Ile Leu Met Leu Ala	
	115 120 125	
60	CGT GTC ATT GTC CAA TTT ACA AAG AAA CAT AAG GAA TGG AGA CGC ACG	432
	Arg Val Ile Val Gln Phe Thr Lys Lys His Lys Glu Trp Arg Arg Thr	
	130 135 140	

-192-

	GTT CCG CTG GAC TTT TTG TTA ATT TTC TGG ATG ATT CTC TTT ACC ATT	480
	Val Pro Leu Asp Phe Leu Leu Ile Phe Trp Met Ile Leu Phe Thr Ile	
	145 150 155 160	
5	CCA GTC CTA GTT CTT TTA GCA CTT CAA AGT GAC TTG GGG ACG GCT TTG	528
	Pro Val Leu Val Leu Leu Ala Leu Gln Ser Asp Leu Gly Thr Ala Leu	
	165 170 175	
10	GTT TTT GTA GCC ATT TTC TCA GGA ATC GTT TTA TTA TCA GGG GTT TCT	576
	Val Phe Val Ala Ile Phe Ser Gly Ile Val Leu Leu Ser Gly Val Ser	
	180 185 190	
15	TGG AAA ATT ATT ATC CCA GTA TTT GTG ACT GCT GTA ACA GGA GTT GCT	624
	Trp Lys Ile Ile Ile Pro Val Phe Val Thr Ala Val Thr Gly Val Ala	
	195 200 205	
20	GGT TTC TTA GCT ATC TTT ATT AGC AAG GAC GGA CGA GCT TTT CTT CAC	672
	Gly Phe Leu Ala Ile Phe Ile Ser Lys Asp Gly Arg Ala Phe Leu His	
	210 215 220	
25	CAG ATT GGA ATG CCG ACC TAC CAA ATC AAT CGG ATT TTG GCT TGG CTC	720
	Gln Ile Gly Met Pro Thr Tyr Gln Ile Asn Arg Ile Leu Ala Trp Leu	
	225 230 235 240	
30	AAT CCC TTT GAG TTT GCC CAA ACA ACG ACT TAC CAG CAG GCT CAA GGG	768
	Asn Pro Phe Glu Phe Ala Gln Thr Thr Thr Tyr Gln Gln Ala Gln Gly	
	245 250 255	
35	CAG ATT GCC ATT GGG AGT GGT GGC TTA TTT GGT CAG GGA TTT AAT GCT	816
	Gln Ile Ala Ile Gly Ser Gly Gly Leu Phe Gly Gln Gly Phe Asn Ala	
	260 265 270	
40	TCG AAT CTG CTT ATC CCA GTT CGA GAG TCA GAT ATG ATT TTT ACG GTT	864
	Ser Asn Leu Leu Ile Pro Val Arg Glu Ser Asp Met Ile Phe Thr Val	
	275 280 285	
45	ATT GCA GAA GAT TTT GGC TTT ATT GGC TCT GTC CTG GTT ATT GCC CTC	912
	Ile Ala Glu Asp Phe Gly Phe Ile Gly Ser Val Leu Val Ile Ala Leu	
	290 295 300	
50	TAT CTC ATG TTG ATT TAC CGT ATG TTG AAG ATT ACT CTT AAA TCA AAT	960
	Tyr Leu Met Leu Ile Tyr Arg Met Leu Lys Ile Thr Leu Lys Ser Asn	
	305 310 315 320	
55	AAC CAG TTC TAC ACT TAT ATT TCC ACA GGT TTG ATT ATG ATG TTG CTC	1008
	Asn Gln Phe Tyr Thr Tyr Ile Ser Thr Gly Leu Ile Met Met Leu Leu	
	325 330 335	
60	TTC CAC ATC TTT GAG AAT ATC GGT GCT GTG ACT GGA CTA CTT CCT TTG	1056
	Phe His Ile Phe Glu Asn Ile Gly Ala Val Thr Gly Leu Leu Pro Leu	
	340 345 350	
65	ACG GGG ATT CCC TTG CCT TTC ATT TCG CAA GGG GGA TCA GCG ATT ATC	1104
	Thr Gly Ile Pro Leu Pro Phe Ile Ser Gln Gly Gly Ser Ala Ile Ile	
	355 360 365	
70	AGT AAT CTG ATT GGT GTT GGT TTG CTT TTA TCG ATG AGT TAC CAG ACT	1152
	Ser Asn Leu Ile Gly Val Gly Leu Leu Leu Ser Met Ser Tyr Gln Thr	
	370 375 380	

-193-

AAT CTA GCT GAA GAA AAG AGC GGA AAA GTC CCA TTC AAA CGG AAA AAG 1200  
 Asn Leu Ala Glu Glu Lys Ser Gly Lys Val Pro Phe Lys Arg Lys Lys  
 385 390 395 400

5 GTT GTA TTA AAA CAA ATT AAA TAA 1224  
 Val Val Leu Lys Gln Ile Lys  
 405

10 (2) INFORMATION FOR SEQ ID NO:114:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 407 amino acids  
 (B) TYPE: amino acid  
 15 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: protein  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

20 Met Lys Arg Ser Leu Asp Ser Arg Val Asp Tyr Ser Leu Leu Leu Pro  
 1 5 10 15  
 Val Phe Phe Leu Leu Val Ile Gly Val Val Ala Ile Tyr Ile Ala Val  
 25 20 25 30  
 Ser His Asp Tyr Pro Asn Asn Ile Leu Pro Ile Leu Gly Gln Gln Val  
 35 35 40 45  
 30 Ala Trp Ile Ala Leu Gly Leu Val Ile Gly Phe Val Val Met Leu Phe  
 50 55 60  
 Asn Thr Glu Phe Leu Trp Lys Val Thr Pro Phe Leu Tyr Ile Leu Gly  
 65 70 75 80  
 35 Leu Gly Leu Met Ile Leu Pro Ile Val Phe Tyr Asn Pro Ser Leu Val  
 85 90 95  
 40 Ala Ser Thr Gly Ala Lys Asn Trp Val Ser Ile Asn Gly Ile Thr Leu  
 100 105 110  
 Phe Gln Pro Ser Glu Phe Met Lys Ile Ser Tyr Ile Leu Met Leu Ala  
 115 120 125  
 45 Arg Val Ile Val Gln Phe Thr Lys Lys His Lys Glu Trp Arg Arg Thr  
 130 135 140  
 Val Pro Leu Asp Phe Leu Leu Ile Phe Trp Met Ile Leu Phe Thr Ile  
 145 150 155 160  
 50 Pro Val Leu Val Leu Leu Ala Leu Gln Ser Asp Leu Gly Thr Ala Leu  
 165 170 175  
 Val Phe Val Ala Ile Phe Ser Gly Ile Val Leu Leu Ser Gly Val Ser  
 55 180 185 190  
 Trp Lys Ile Ile Ile Pro Val Phe Val Thr Ala Val Thr Gly Val Ala  
 195 200 205  
 60 Gly Phe Leu Ala Ile Phe Ile Ser Lys Asp Gly Arg Ala Phe Leu His  
 210 215 220

-194-

Gln Ile Gly Met Pro Thr Tyr Gln Ile Asn Arg Ile Leu Ala Trp Leu  
 225 230 235 240  
 5 Asn Pro Phe Glu Phe Ala Gln Thr Thr Thr Tyr Gln Gln Ala Gln Gly  
 245 250 255  
 Gln Ile Ala Ile Gly Ser Gly Gly Leu Phe Gly Gln Gly Phe Asn Ala  
 260 265 270  
 10 Ser Asn Leu Leu Ile Pro Val Arg Glu Ser Asp Met Ile Phe Thr Val  
 275 280 285  
 Ile Ala Glu Asp Phe Gly Phe Ile Gly Ser Val Leu Val Ile Ala Leu  
 290 295 300  
 15 Tyr Leu Met Leu Ile Tyr Arg Met Leu Lys Ile Thr Leu Lys Ser Asn  
 305 310 315 320  
 20 Asn Gln Phe Tyr Thr Tyr Ile Ser Thr Gly Leu Ile Met Met Leu Leu  
 325 330 335  
 Phe His Ile Phe Glu Asn Ile Gly Ala Val Thr Gly Leu Leu Pro Leu  
 340 345 350  
 25 Thr Gly Ile Pro Leu Pro Phe Ile Ser Gln Gly Gly Ser Ala Ile Ile  
 355 360 365  
 Ser Asn Leu Ile Gly Val Gly Leu Leu Leu Ser Met Ser Tyr Gln Thr  
 370 375 380  
 30 Asn Leu Ala Glu Glu Lys Ser Gly Lys Val Pro Phe Lys Arg Lys Lys  
 385 390 395 400  
 35 Val Val Leu Lys Gln Ile Lys  
 405

## (2) INFORMATION FOR SEQ ID NO:115:

- 40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1311 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 45 (ii) MOLECULE TYPE: DNA (genomic)  
 (iii) HYPOTHETICAL: NO  
 50 (iv) ANTI-SENSE: NO  
 (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 55 (B) LOCATION: 1..1311  
 (D) OTHER INFORMATION: SecY

- 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

ATG TTT TTT AAA TTA TTA AGA GAA GCT CTT AAA GTC AAG CAG GTT CGA

-195-

	Met	Phe	Phe	Lys	Leu	Leu	Arg	Glu	Ala	Leu	Lys	Val	Lys	Gln	Val	Arg	
	1				5					10					15		
5	TCA	AAA	ATT	TTA	TTT	ACA	ATT	TTT	ATC	GTT	TTG	GTC	TTT	CGT	ATC	GGA	96
	Ser	Lys	Ile	Leu	Phe	Thr	Ile	Phe	Ile	Val	Leu	Val	Phe	Arg	Ile	Gly	
				20					25					30			
10	ACT	AGC	ATT	ACA	GTT	CCT	GGT	GTG	AAT	GCC	AAT	AGC	TTG	AAT	GCT	TTA	144
	Thr	Ser	Ile	Thr	Val	Pro	Gly	Val	Asn	Ala	Asn	Ser	Leu	Asn	Ala	Leu	
				35				40					45				
15	AGT	GGA	TTA	TCC	TTC	TTA	AAC	ATG	TTG	AGC	TTG	GTG	TCG	GGG	AAT	GCC	192
	Ser	Gly	Leu	Ser	Phe	Leu	Asn	Met	Leu	Ser	Leu	Val	Ser	Gly	Asn	Ala	
		50					55					60					
20	CTA	AAA	AAC	TTT	TCG	ATT	TTT	GCC	CTA	GGA	GTT	AGT	CCC	TAT	ATC	ACC	240
	Leu	Lys	Asn	Phe	Ser	Ile	Phe	Ala	Leu	Gly	Val	Ser	Pro	Tyr	Ile	Thr	
	65					70					75					80	
25	GCT	TCT	ATT	GTT	GTC	CAA	CTC	TTG	CAA	ATG	GAT	ATT	TTA	CCC	AAG	TTT	288
	Ala	Ser	Ile	Val	Val	Gln	Leu	Leu	Gln	Met	Asp	Ile	Leu	Pro	Lys	Phe	
					85				90						95		
30	GTA	GAG	TGG	GGT	AAA	CAA	GGG	GAA	GTA	GGT	CGA	AGA	AAA	TTG	AAT	CAA	336
	Val	Glu	Trp	Gly	Lys	Gln	Gly	Glu	Val	Gly	Arg	Arg	Lys	Leu	Asn	Gln	
			100					105						110			
35	GCT	ACT	CGT	TAT	ATT	GCT	CTA	GTT	CTC	GCT	TTT	GTG	CAA	TCT	ATC	GGG	384
	Ala	Thr	Arg	Tyr	Ile	Ala	Leu	Val	Leu	Ala	Phe	Val	Gln	Ser	Ile	Gly	
			115					120					125				
40	ATT	ACA	GCT	GGT	TTT	AAT	ACC	TTG	GCT	GGA	GCT	CAA	TTG	ATT	AAA	ACT	432
	Ile	Thr	Ala	Gly	Phe	Asn	Thr	Leu	Ala	Gly	Ala	Gln	Leu	Ile	Lys	Thr	
			130				135					140					
45	GCT	TTA	ACT	CCA	CAA	GTT	TTT	CTG	ACG	ATT	GGT	ATC	ATC	TTA	ACA	GCT	480
	Ala	Leu	Thr	Pro	Gln	Val	Phe	Leu	Thr	Ile	Gly	Ile	Ile	Leu	Thr	Ala	
	145					150					155					160	
50	GGT	AGT	ATG	ATT	GTC	ACT	TGG	TTG	GGT	GAG	CAA	ATT	ACA	GAT	AAG	GGA	528
	Gly	Ser	Met	Ile	Val	Thr	Trp	Leu	Gly	Glu	Gln	Ile	Thr	Asp	Lys	Gly	
					165				170						175		
55	TAC	GGA	AAC	GGT	GTT	TCC	ATG	ATT	ATC	TTT	GCC	GGG	ATT	GTT	TCC	TCA	576
	Tyr	Gly	Asn	Gly	Val	Ser	Met	Ile	Ile	Phe	Ala	Gly	Ile	Val	Ser	Ser	
				180					185					190			
60	ATT	CCA	GAG	ATG	ATT	CAG	GGC	ATC	TAT	GTG	GAC	TAC	TTT	GTG	AAC	GTC	624
	Ile	Pro	Glu	Met	Ile	Gln	Gly	Ile	Tyr	Val	Asp	Tyr	Phe	Val	Asn	Val	
			195					200					205				
65	CCA	AGT	AGC	CGT	ATC	ACT	TCA	TCT	ATC	ATT	TTC	GTA	ATC	ATT	TTG	ATT	672
	Pro	Ser	Ser	Arg	Ile	Thr	Ser	Ser	Ile	Ile	Phe	Val	Ile	Ile	Leu	Ile	
			210					215					220				
70	ATT	ACT	GTA	TTG	TTG	ATT	ATT	TAC	TTT	ACA	ACT	TAT	GTT	CAA	CAA	GCA	720
	Ile	Thr	Val	Leu	Leu	Ile	Ile	Tyr	Phe	Thr	Thr	Tyr	Val	Gln	Gln	Ala	
	225					230					235					240	
75	GAA	TAC	AAA	ATT	CCA	ATC	CAA	TAT	ACT	AAG	GTT	GCA	CAA	GGT	GCT	CCA	768
	Glu	Tyr	Lys	Ile	Pro	Ile	Gln	Tyr	Thr	Lys	Val	Ala	Gln	Gly	Ala	Pro	



-196-

	245	250	255	
5	TCT AGC TCT TAC CTT CCG TTA AAG GTA AAT CCT GCT GGA GTT ATC CCT Ser Ser Ser Tyr Leu Pro Leu Lys Val Asn Pro Ala Gly Val Ile Pro 260 265 270			816
10	GTT ATC TTT GCC AGT TCG ATT ACT GCA GCG CCT GCG GCT ATT CTT CAG Val Ile Phe Ala Ser Ser Ile Thr Ala Ala Pro Ala Ala Ile Leu Gln 275 280 285			864
15	TTT TTG AGT GCC ACA GGT CAT GAT TGG GCT TGG GTA AGG GTA GCA CAA Phe Leu Ser Ala Thr Gly His Asp Trp Ala Trp Val Arg Val Ala Gln 290 295 300			912
20	GAG ATG TTG GCA ACT ACT TCT CCA ACT GGT ATT GCC ATG TAT GCT TTG Glu Met Leu Ala Thr Thr Ser Pro Thr Gly Ile Ala Met Tyr Ala Leu 305 310 315 320			960
25	TTG ATT ATT CTC TTT ACA TTC TTC TAT ACG TTT GTA CAG ATT AAT CCT Leu Ile Ile Leu Phe Thr Phe Phe Tyr Thr Phe Val Gln Ile Asn Pro 325 330 335			1008
30	GAA AAA GCA GCA GAG AGC CTA CAA AAG AGT GGT GCC TAT ATC CAT GGA Glu Lys Ala Ala Glu Ser Leu Gln Lys Ser Gly Ala Tyr Ile His Gly 340 345 350			1056
35	GTT CGT CCT GGT AAA GGT ACA GAA GAA TAT ATG TCT AAA CTT CTT CGT Val Arg Pro Gly Lys Gly Thr Glu Glu Tyr Met Ser Lys Leu Leu Arg 355 360 365			1104
40	CGT CTT GCA ACT GTT GGT TCC CTC TTC CTT GGT GTG ATT TCC ATT TTA Arg Leu Ala Thr Val Gly Ser Leu Phe Leu Gly Val Ile Ser Ile Leu 370 375 380			1152
45	CCG ATT GCA GCT AAA GAT GTA TTT GGT CTT TCT GAT GTT GTT GCC TTT Pro Ile Ala Ala Lys Asp Val Phe Gly Leu Ser Asp Val Val Ala Phe 385 390 395 400			1200
50	GGT GGA ACA AGT CTC TTG ATC ATT ATC TCT ACA GGT ATC GAA GGA ATC Gly Gly Thr Ser Leu Ile Ile Ile Ser Thr Gly Ile Glu Gly Ile 405 410 415			1248
55	AAG CAA TTG GAA GGT TAC CTA TTG AAA CGT AAG TAT GTT GGT TTC ATG Lys Gln Leu Glu Gly Tyr Leu Leu Lys Arg Lys Tyr Val Gly Phe Met 420 425 430			1296
60	GAC AGA ACA GAA TAA Asp Arg Thr Glu * 435			1311

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 437 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

-197-

Met Phe Phe Lys Leu Leu Arg Glu Ala Leu Lys Val Lys Gln Val Arg  
 1 5 10 15  
 5 Ser Lys Ile Leu Phe Thr Ile Phe Ile Val Leu Val Phe Arg Ile Gly  
 20 25 30  
 Thr Ser Ile Thr Val Pro Gly Val Asn Ala Asn Ser Leu Asn Ala Leu  
 35 40 45  
 10 Ser Gly Leu Ser Phe Leu Asn Met Leu Ser Leu Val Ser Gly Asn Ala  
 50 55 60  
 Leu Lys Asn Phe Ser Ile Phe Ala Leu Gly Val Ser Pro Tyr Ile Thr  
 65 70 75 80  
 15 Ala Ser Ile Val Val Gln Leu Leu Gln Met Asp Ile Leu Pro Lys Phe  
 85 90 95  
 20 Val Glu Trp Gly Lys Gln Gly Glu Val Gly Arg Arg Lys Leu Asn Gln  
 100 105 110  
 Ala Thr Arg Tyr Ile Ala Leu Val Leu Ala Phe Val Gln Ser Ile Gly  
 115 120 125  
 25 Ile Thr Ala Gly Phe Asn Thr Leu Ala Gly Ala Gln Leu Ile Lys Thr  
 130 135 140  
 Ala Leu Thr Pro Gln Val Phe Leu Thr Ile Gly Ile Ile Leu Thr Ala  
 145 150 155 160  
 30 Gly Ser Met Ile Val Thr Trp Leu Gly Glu Gln Ile Thr Asp Lys Gly  
 165 170 175  
 35 Tyr Gly Asn Gly Val Ser Met Ile Ile Phe Ala Gly Ile Val Ser Ser  
 180 185 190  
 Ile Pro Glu Met Ile Gln Gly Ile Tyr Val Asp Tyr Phe Val Asn Val  
 195 200 205  
 40 Pro Ser Ser Arg Ile Thr Ser Ser Ile Ile Phe Val Ile Ile Leu Ile  
 210 215 220  
 Ile Thr Val Leu Leu Ile Ile Tyr Phe Thr Thr Tyr Val Gln Gln Ala  
 225 230 235 240  
 45 Glu Tyr Lys Ile Pro Ile Gln Tyr Thr Lys Val Ala Gln Gly Ala Pro  
 245 250 255  
 50 Ser Ser Ser Tyr Leu Pro Leu Lys Val Asn Pro Ala Gly Val Ile Pro  
 260 265 270  
 Val Ile Phe Ala Ser Ser Ile Thr Ala Ala Pro Ala Ala Ile Leu Gln  
 275 280 285  
 55 Phe Leu Ser Ala Thr Gly His Asp Trp Ala Trp Val Arg Val Ala Gln  
 290 295 300  
 Glu Met Leu Ala Thr Thr Ser Pro Thr Gly Ile Ala Met Tyr Ala Leu  
 305 310 315 320  
 60

-198-

Leu Ile Ile Leu Phe Thr Phe Phe Tyr Thr Phe Val Gln Ile Asn Pro  
                             325                            330                            335  
 5   Glu Lys Ala Ala Glu Ser Leu Gln Lys Ser Gly Ala Tyr Ile His Gly  
                             340                            345                            350  
   Val Arg Pro Gly Lys Gly Thr Glu Glu Tyr Met Ser Lys Leu Leu Arg  
                             355                            360                            365  
 10   Arg Leu Ala Thr Val Gly Ser Leu Phe Leu Gly Val Ile Ser Ile Leu  
                             370                            375                            380  
   Pro Ile Ala Ala Lys Asp Val Phe Gly Leu Ser Asp Val Val Ala Phe  
                             385                            390                            395                            400  
 15   Gly Gly Thr Ser Leu Leu Ile Ile Ile Ser Thr Gly Ile Glu Gly Ile  
                             405                            410                            415  
 20   Lys Gln Leu Glu Gly Tyr Leu Leu Lys Arg Lys Tyr Val Gly Phe Met  
                             420                            425                            430  
   Asp Arg Thr Glu  
                             435

25   (2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:  
       (A) LENGTH: 1959 base pairs  
       (B) TYPE: nucleic acid  
 30       (C) STRANDEDNESS: single  
       (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

35   (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40   (ix) FEATURE:

      (A) NAME/KEY: CDS  
       (B) LOCATION: 1..1956  
       (D) OTHER INFORMATION: FtsH

45   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

50	ATG AAA AAA CAA AAT AAT GGT TTA ATT AAA AAT CCT TTT CTA TGG TTA Met Lys Lys Gln Asn Asn Gly Leu Ile Lys Asn Pro Phe Leu Trp Leu 1                            5                            10                            15	48
55	TTA TTT ATC TTT TTC CTT GTG ACA GGA TTC CAG TAT TTC TAT TCT GGG Leu Phe Ile Phe Phe Leu Val Thr Gly Phe Gln Tyr Phe Tyr Ser Gly 20                            25                            30	96
60	AAT AAC TCA GGA GGA AGT CAG CAA ATC AAC TAT ACT GAG TTG GTA CAA Asn Asn Ser Gly Gly Ser Gln Ile Asn Tyr Thr Glu Leu Val Gln 35                            40                            45	144
60	GAA ATT ACC GAT GGT AAT GAA AAA GAA TTA ACT TAC CAA CCA AAT GTT Glu Ile Thr Asp Gly Asn Glu Lys Glu Leu Thr Tyr Gln Pro Asn Val	192

-199-

	50		55		60	
5	AGT GTT ATC GAA GTT TCT GGT GTC TAT AAA AAT CCT AAA ACA AGT AAA Ser Val Ile Glu Val Ser Gly Val Tyr Lys Asn Pro Lys Thr Ser Lys 65 70 75 80	240				
10	GAA GGA ACA GGT ATT CAG TTT TTC ACG CCA TCT GTT ACT AAG GTA GAG Glu Gly Thr Gly Ile Gln Phe Phe Thr Pro Ser Val Thr Lys Val Glu 85 90 95	288				
15	AAA TTT ACC AGC ACT ATT CTT CCT GCA GAT ACT ACC GTA TCA GAA TTG Lys Phe Thr Ser Thr Ile Leu Pro Ala Asp Thr Thr Val Ser Glu Leu 100 105 110	336				
20	CAA AAA CTT GCT ACT GAC CAT AAA GCA GAA GTA ACT GTT AAG CAT GAA Gln Lys Leu Ala Thr Asp His Lys Ala Glu Val Thr Val Lys His Glu 115 120 125	384				
25	AGT TCA AGT GGT ATA TGG ATT AAT CTA CTC GTA TCC ATT GTG CCA TTT Ser Ser Ser Gly Ile Trp Ile Asn Leu Leu Val Ser Ile Val Pro Phe 130 135 140	432				
30	GGA ATT CTA TTC TTC TTC CTA TTC TCT ATG ATG GGA AAT ATG GGA GGA Gly Ile Leu Phe Phe Phe Leu Phe Ser Met Met Gly Asn Met Gly Gly 145 150 155 160	480				
35	GGC AAT GGC CGT AAT CCA ATG AGT TTT GGA CGT AGT AAG GCT AAA GCA Gly Asn Gly Arg Asn Pro Met Ser Phe Gly Arg Ser Lys Ala Lys Ala 165 170 175	528				
40	GCA AAT AAA GAA GAT ATT AAA GTA AGA TTT TCA GAT GTT GCT GGA GCT Ala Asn Lys Glu Asp Ile Lys Val Arg Phe Ser Asp Val Ala Gly Ala 180 185 190	576				
45	GAG GAA GAA AAA CAA GAA CTA GTT GAA GTT GTT GAG TTC TTA AAA GAT Glu Glu Glu Lys Gln Glu Leu Val Glu Val Val Glu Phe Leu Lys Asp 195 200 205	624				
50	CCA AAA CGA TTC ACA AAA CTT GGA GCC CGT ATT CCA GCA GGT GTT CTT Pro Lys Arg Phe Thr Lys Leu Gly Ala Arg Ile Pro Ala Gly Val Leu 210 215 220	672				
55	TTG GAG GGA CCT CCG GGG ACA GGT AAG ACT TTG CTT GCT AAG GCA GTC Leu Glu Gly Pro Pro Gly Thr Gly Lys Thr Leu Leu Ala Lys Ala Val 225 230 235 240	720				
60	GCT GGA GAA GCA GGT GTT CCA TTC TTT AGT ATC TCA GGT TCT GAC TTT Ala Gly Glu Ala Gly Val Pro Phe Phe Ser Ile Ser Gly Ser Asp Phe 245 250 255	768				
65	GTA GAA ATG TTT GTC GGA GTT GGA GCT AGT CGT GTT CGC TCT CTT TTT Val Glu Met Phe Val Gly Val Gly Ala Ser Arg Val Arg Ser Leu Phe 260 265 270	816				
70	GAG GAT GCC AAA AAA GCA GCA CCA GCT ATC ATC TTT ATC GAT CTA AAT Glu Asp Ala Lys Lys Ala Ala Pro Ala Ile Ile Phe Ile Asp Leu Asn 275 280 285	864				
75	GAT GCT GTT GGA CGT CAA CGT GGA GTC GGT CTC GGC GGA GGT AAT GAC Asp Ala Val Gly Arg Gln Arg Gly Val Gly Leu Gly Gly Gly Asn Asp 290 295 300	912				

-200-

	GAA CGT GAA CAA ACC TTG AAC CAA CTT TTG ATT GAG ATG GAT GGT TTT	960
	Glu Arg Glu Gln Thr Leu Asn Gln Leu Leu Ile Glu Met Asp Gly Phe	
	305 310 315 320	
5	GAG GGA AAT GAA GGG ATT ATC GTC ATC GCT GCG ACA AAC CGT TCA GAT	1008
	Glu Gly Asn Glu Gly Ile Ile Val Ile Ala Ala Thr Asn Arg Ser Asp	
	325 330 335	
10	GTA CTT GAT CCT GCC CTT TTG CGT CCA GGA CGT TTT GAT AGA AAA GTA	1056
	Val Leu Asp Pro Ala Leu Leu Arg Pro Gly Arg Phe Asp Arg Lys Val	
	340 345 350	
15	TTG GTT GGC CGT CCT GAT GTT AAA GGT CGT GAA GCA ATC TTG AAA GTT	1104
	Leu Val Gly Arg Pro Asp Val Lys Gly Arg Glu Ala Ile Leu Lys Val	
	355 360 365	
20	CAC GCT AAG AAC AAG CCT TTA GCA GAA GAT GTT GAT TTG AAA TTA GTG	1152
	His Ala Lys Asn Lys Pro Leu Ala Glu Asp Val Asp Leu Lys Leu Val	
	370 375 380	
25	GCT CAA CAA ACT CCA GGC TTT GTT GGT GCT GAT TTA GAG AAT GTC TTG	1200
	Ala Gln Gln Thr Pro Gly Phe Val Gly Ala Asp Leu Glu Asn Val Leu	
	385 390 395 400	
30	AAT GAA GCA GCT TTA GTT GCT GCT CGT CGC AAT AAA TCG ATA ATT GAT	1248
	Asn Glu Ala Ala Leu Val Ala Ala Arg Arg Asn Lys Ser Ile Ile Asp	
	405 410 415	
35	GCT TCA GAT ATT GAT GAA GCA GAA GAT AGA GTT ATT GCT GGA CCT TCT	1296
	Ala Ser Asp Ile Asp Glu Ala Glu Asp Arg Val Ile Ala Gly Pro Ser	
	420 425 430	
40	AAG AAA GAT AAG ACA GTT TCA CAA AAA GAA CGA GAA TTG GTT GCT TAC	1344
	Lys Lys Asp Lys Thr Val Ser Gln Lys Glu Arg Glu Leu Val Ala Tyr	
	435 440 445	
45	CAT GAG GCA GGA CAT ACC ATT GTT GGT CTA GTC TTG TCG ACT GCT CGC	1392
	His Glu Ala Gly His Thr Ile Val Gly Leu Val Leu Ser Thr Ala Arg	
	450 455 460	
50	GTT GTC CAT AAG GTT ACA ATT GTA CCA CGC GGC CGT GCA GGC GGA TAC	1440
	Val Val His Lys Val Thr Ile Val Pro Arg Gly Arg Ala Gly Gly Tyr	
	465 470 475 480	
55	ATG ATT GCA CTT CCT AAA GAG GAT CAA ATG CTT CTA TCT AAA GAA GAT	1488
	Met Ile Ala Leu Pro Lys Glu Asp Gln Met Leu Leu Ser Lys Glu Asp	
	485 490 495	
60	ATG AAA GAG CAA TTG GCT GGC TTA ATG GGT GGA CGT GTA GCT GAA GAA	1536
	Met Lys Glu Gln Leu Ala Gly Leu Met Gly Gly Arg Val Ala Glu Glu	
	500 505 510	
65	ATT ATC TTT AAT GTC CAA ACT ACA GGA GCT TCA AAC GAC TTT GAA CAA	1584
	Ile Ile Phe Asn Val Gln Thr Thr Gly Ala Ser Asn Asp Phe Glu Gln	
	515 520 525	
70	GCG ACA CAA ATG GCA CGT GCA ATG GTT ACA GAG TAC GGT ATG AGT GAA	1632
	Ala Thr Gln Met Ala Arg Ala Met Val Thr Glu Tyr Gly Met Ser Glu	
	530 535 540	

-201-

AAA CTT GGC CCA GTA CAA TAT GAA GGA AAC CAT GCT ATG CTT GGT GCA 1680  
 Lys Leu Gly Pro Val Gln Tyr Glu Gly Asn His Ala Met Leu Gly Ala  
 545 550 555 560

5 CAG AGT CCT CAA AAA TCA ATT TCA GAA CAA ACA GCT TAT GAA ATT GAT 1728  
 Gln Ser Pro Gln Lys Ser Ile Ser Glu Gln Thr Ala Tyr Glu Ile Asp  
 565 570 575

10 GAA GAG GTT CGT TCA TTA TTA AAT GAG GCA CGA AAT AAA GCT GCT GAA 1776  
 Glu Glu Val Arg Ser Leu Leu Asn Glu Ala Arg Asn Lys Ala Ala Glu  
 580 585 590

15 ATT ATT CAG TCA AAT CGT GAA ACT CAC AAG TTA ATT GCA GAA GCA TTA 1824  
 Ile Ile Gln Ser Asn Arg Glu Thr His Lys Leu Ile Ala Glu Ala Leu  
 595 600 605

20 TTG AAA TAC GAA ACA TTG GAT AGT ACA CAA ATT AAA GCT CTT TAC GAA 1872  
 Leu Lys Tyr Glu Thr Leu Asp Ser Thr Gln Ile Lys Ala Leu Tyr Glu  
 610 615 620

ACA GGA AAG ATG CCT GAA GCA GTA GAA GAG GAA TCT CAT GCA CTA TCC 1920  
 Thr Gly Lys Met Pro Glu Ala Val Glu Glu Glu Ser His Ala Leu Ser  
 625 630 635 640

25 TAT GAT GAA GTA AAG TCA AAA ATG AAT GAC GAA AAA TAA 1959  
 Tyr Asp Glu Val Lys Ser Lys Met Asn Asp Glu Lys  
 645 650

30 (2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 652 amino acids  
 (B) TYPE: amino acid  
 35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

40 Met Lys Lys Gln Asn Asn Gly Leu Ile Lys Asn Pro Phe Leu Trp Leu  
 1 5 10 15

45 Leu Phe Ile Phe Phe Leu Val Thr Gly Phe Gln Tyr Phe Tyr Ser Gly  
 20 25 30

Asn Asn Ser Gly Gly Ser Gln Gln Ile Asn Tyr Thr Glu Leu Val Gln  
 35 40 45

50 Glu Ile Thr Asp Gly Asn Glu Lys Glu Leu Thr Tyr Gln Pro Asn Val  
 50 55 60

Ser Val Ile Glu Val Ser Gly Val Tyr Lys Asn Pro Lys Thr Ser Lys  
 65 70 75 80

55 Glu Gly Thr Gly Ile Gln Phe Phe Thr Pro Ser Val Thr Lys Val Glu  
 85 90 95

60 Lys Phe Thr Ser Thr Ile Leu Pro Ala Asp Thr Thr Val Ser Glu Leu  
 100 105 110

-202-

Gln Lys Leu Ala Thr Asp His Lys Ala Glu Val Thr Val Lys His Glu  
 115 120 125  
 5 Ser Ser Ser Gly Ile Trp Ile Asn Leu Leu Val Ser Ile Val Pro Phe  
 130 135 140  
 Gly Ile Leu Phe Phe Phe Leu Phe Ser Met Met Gly Asn Met Gly Gly  
 145 150 155 160  
 10 Gly Asn Gly Arg Asn Pro Met Ser Phe Gly Arg Ser Lys Ala Lys Ala  
 165 170 175  
 Ala Asn Lys Glu Asp Ile Lys Val Arg Phe Ser Asp Val Ala Gly Ala  
 180 185 190  
 15 Glu Glu Glu Lys Gln Glu Leu Val Glu Val Val Glu Phe Leu Lys Asp  
 195 200 205  
 Pro Lys Arg Phe Thr Lys Leu Gly Ala Arg Ile Pro Ala Gly Val Leu  
 210 215 220  
 20 Leu Glu Gly Pro Pro Gly Thr Gly Lys Thr Leu Leu Ala Lys Ala Val  
 225 230 235 240  
 25 Ala Gly Glu Ala Gly Val Pro Phe Phe Ser Ile Ser Gly Ser Asp Phe  
 245 250 255  
 Val Glu Met Phe Val Gly Val Gly Ala Ser Arg Val Arg Ser Leu Phe  
 260 265 270  
 30 Glu Asp Ala Lys Lys Ala Ala Pro Ala Ile Ile Phe Ile Asp Leu Asn  
 275 280 285  
 35 Asp Ala Val Gly Arg Gln Arg Gly Val Gly Leu Gly Gly Asn Asp  
 290 295 300  
 Glu Arg Glu Gln Thr Leu Asn Gln Leu Leu Ile Glu Met Asp Gly Phe  
 305 310 315 320  
 40 Glu Gly Asn Glu Gly Ile Ile Val Ile Ala Ala Thr Asn Arg Ser Asp  
 325 330 335  
 Val Leu Asp Pro Ala Leu Leu Arg Pro Gly Arg Phe Asp Arg Lys Val  
 340 345 350  
 45 Leu Val Gly Arg Pro Asp Val Lys Gly Arg Glu Ala Ile Leu Lys Val  
 355 360 365  
 His Ala Lys Asn Lys Pro Leu Ala Glu Asp Val Asp Leu Lys Leu Val  
 370 375 380  
 50 Ala Gln Gln Thr Pro Gly Phe Val Gly Ala Asp Leu Glu Asn Val Leu  
 385 390 395 400  
 55 Asn Glu Ala Ala Leu Val Ala Ala Arg Arg Asn Lys Ser Ile Ile Asp  
 405 410 415  
 Ala Ser Asp Ile Asp Glu Ala Glu Asp Arg Val Ile Ala Gly Pro Ser  
 420 425 430  
 60 Lys Lys Asp Lys Thr Val Ser Gln Lys Glu Arg Glu Leu Val Ala Tyr

-203-

	435	440	445
	His Glu Ala Gly His Thr Ile Val Gly Leu Val Leu Ser Thr Ala Arg		
	450	455	460
5	Val Val His Lys Val Thr Ile Val Pro Arg Gly Arg Ala Gly Gly Tyr		
	465	470	475
	Met Ile Ala Leu Pro Lys Glu Asp Gln Met Leu Leu Ser Lys Glu Asp		
10		485	490
	Met Lys Glu Gln Leu Ala Gly Leu Met Gly Gly Arg Val Ala Glu Glu		
		500	505
	Ile Ile Phe Asn Val Gln Thr Thr Gly Ala Ser Asn Asp Phe Glu Gln		
15		515	520
	Ala Thr Gln Met Ala Arg Ala Met Val Thr Glu Tyr Gly Met Ser Glu		
20		530	535
	Lys Leu Gly Pro Val Gln Tyr Glu Gly Asn His Ala Met Leu Gly Ala		
	545	550	555
	Gln Ser Pro Gln Lys Ser Ile Ser Glu Gln Thr Ala Tyr Glu Ile Asp		
25		565	570
	Glu Glu Val Arg Ser Leu Leu Asn Glu Ala Arg Asn Lys Ala Ala Glu		
		580	585
	Ile Ile Gln Ser Asn Arg Glu Thr His Lys Leu Ile Ala Glu Ala Leu		
30		595	600
	Leu Lys Tyr Glu Thr Leu Asp Ser Thr Gln Ile Lys Ala Leu Tyr Glu		
35		610	615
	Thr Gly Lys Met Pro Glu Ala Val Glu Glu Glu Ser His Ala Leu Ser		
	625	630	635
	Tyr Asp Glu Val Lys Ser Lys Met Asn Asp Glu Lys		
40		645	650

## (2) INFORMATION FOR SEQ ID NO:119:

- 45 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1278 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- 50 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 55 (iv) ANTI-SENSE: NO
- (ix) FEATURE:
- (A) NAME/KEY: CDS
  - (B) LOCATION: 1..1278
  - (D) OTHER INFORMATION: FtsY
- 60



-204-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

5	ATG GGA TTG TTT GAC CGT CTA TTC GGA AAA AAA GAA GAA CCT AAA ATC	48
	Met Gly Leu Phe Asp Arg Leu Phe Gly Lys Lys Glu Glu Pro Lys Ile	
	1 5 10 15	
10	GAA GAA GTT GTA AAA GAA GCT CTG GAA AAT CTT GAT TTG TCT GAA GAT	96
	Glu Glu Val Val Lys Glu Ala Leu Glu Asn Leu Asp Leu Ser Glu Asp	
	20 25 30	
15	GTT GAT CCT ACC TTC ACA GAA GTT GAG GAA GTT TCT CAG GAA GAA GCA	144
	Val Asp Pro Thr Phe Thr Glu Val Glu Glu Val Ser Gln Glu Glu Ala	
	35 40 45	
20	GAG GTT GAA ATT GTT GAA CAA GCT GTG TTC CAA GAA GAG GAA ATC CAA	192
	Glu Val Glu Ile Val Glu Gln Ala Val Phe Gln Glu Glu Glu Ile Gln	
	50 55 60	
25	GAC ACA GTT GAA GAA AGT CTG GAT TTA GAG CCA GTT GTA GAA GTT TCT	240
	Asp Thr Val Glu Glu Ser Leu Asp Leu Glu Pro Val Val Glu Val Ser	
	65 70 75 80	
30	CAA AAA GAA GTC GAA GAA TTT CCA CAC TCA GAA GAA GGG AAT ACT GAG	288
	Gln Lys Glu Val Glu Glu Phe Pro His Ser Glu Glu Gly Asn Thr Glu	
	85 90 95	
35	TTT CTA GAG ACT ATA GAA GAA AAT AAT TCT GAA GTT CTT GAA CCA GAA	336
	Phe Leu Glu Thr Ile Glu Glu Asn Asn Ser Glu Val Leu Glu Pro Glu	
	100 105 110	
40	AGG CCT CAA GCA GAA GAA ACC GTT CAG GAA AAA TAT GAC CGC AGT CTT	384
	Arg Pro Gln Ala Glu Glu Thr Val Gln Glu Lys Tyr Asp Arg Ser Leu	
	115 120 125	
45	AAG AAA ACT CGT ACA GGT TTC GGT GCC CGC TTG AAT GCC TTC TTT GCT	432
	Lys Lys Thr Arg Thr Gly Phe Gly Ala Arg Leu Asn Ala Phe Phe Ala	
	130 135 140	
50	AAC TTC CGC TCT GTT GAC GAA GAA TTT TTC GAG GAA CTG GAA GAA CTG	480
	Asn Phe Arg Ser Val Asp Glu Glu Phe Phe Glu Glu Leu Glu Leu	
	145 150 155 160	
55	CTG ATT ATG AGT GAT GTT GGT GTC CAA GTC GCT TCT AAC TTA ACG GAG	528
	Leu Ile Met Ser Asp Val Gly Val Gln Val Ala Ser Asn Leu Thr Glu	
	165 170 175	
60	GAA CTA CGT TAC GAA GCC AAG CTT GAA AAT GCC AAG AAA CCT GAT GCA	576
	Glu Leu Arg Tyr Glu Ala Lys Leu Glu Asn Ala Lys Lys Pro Asp Ala	
	180 185 190	
65	CTT CGT CGT GTC ATC ATT GAG AAA TTG GTT GAG CTT TAT GAA AAG GAT	624
	Leu Arg Arg Val Ile Ile Glu Lys Leu Val Glu Leu Tyr Glu Lys Asp	
	195 200 205	
70	GGT AGC TAC GAT GAA AGC ATC CAC TTC CAA GAT AAC TTG ACA GTT ATG	672
	Gly Ser Tyr Asp Glu Ser Ile His Phe Gln Asp Asn Leu Thr Val Met	
	210 215 220	
75	CTC TTT GTT GGT GTG AAT GGT GTT GGG AAA ACA ACT TCT ATC GGA AAA	720
	Leu Phe Val Gly Val Asn Gly Val Gly Lys Thr Thr Ser Ile Gly Lys	
	225 230 235 240	

-205-

CTA GCC CAC CGC TAC AAA CAA GCT GGT AAG AAG GTC ATG CTG GTT GCA 768  
 Leu Ala His Arg Tyr Lys Gln Ala Gly Lys Lys Val Met Leu Val Ala  
 245 250 255  
 5 GCA GAT ACC TTC CGT GCG GGT GCA GTA GCT CAG CTA GCT GAA TGG GGC 816  
 Ala Asp Thr Phe Arg Ala Gly Ala Val Ala Gln Leu Ala Glu Trp Gly  
 260 265 270  
 10 CGA CGA GTA GAT GTT CCA GTA GTA ACT GGA CCT GAA AAA GCT GAT CCA 864  
 Arg Arg Val Asp Val Pro Val Thr Gly Pro Glu Lys Ala Asp Pro  
 275 280 285  
 15 GCC AGC GTG GTC TTT GAT GGT ATG GAA CGT GCC GTG GCT GAA GGT ATC 912  
 Ala Ser Val Val Phe Asp Gly Met Glu Arg Ala Val Ala Glu Gly Ile  
 290 295 300  
 GAT ATT CTC ATG ATT GAT ACT GCT GGT CGT CTG CAA AAT AAG GAT AAC 960  
 Asp Ile Leu Met Ile Asp Thr Ala Gly Arg Leu Gln Asn Lys Asp Asn  
 305 310 315 320  
 20 CTT ATG GCT GAG TTG GAA AAG ATT GGT CGT ATT ATC AAA CGT GTT GTG 1008  
 Leu Met Ala Glu Leu Glu Lys Ile Gly Arg Ile Ile Lys Arg Val Val  
 325 330 335  
 25 CCA GAA GCA CCA CAT GAA ACC TTC TTG GCA CTT GAT GCA TCA ACA GGT 1056  
 Pro Glu Ala Pro His Glu Thr Phe Leu Ala Leu Asp Ala Ser Thr Gly  
 340 345 350  
 30 CAA AAT GCC CTA GTA CAG GCC AAA GAA TTT TCG AAA ATC ACA CCT TTA 1104  
 Gln Asn Ala Leu Val Gln Ala Lys Glu Phe Ser Lys Ile Thr Pro Leu  
 355 360 365  
 35 ACG GGA ATT GTT TTG ACT AAG ATT GAT GGA ACT GCT CGA GGA GGT GTG 1152  
 Thr Gly Ile Val Leu Thr Lys Ile Asp Gly Thr Ala Arg Gly Gly Val  
 370 375 380  
 GTT CTA GCC ATT CGT GAA GAA CTC AAT ATT CCT GTA AAA TTG ATT GGT 1200  
 Val Leu Ala Ile Arg Glu Glu Leu Asn Ile Pro Val Lys Leu Ile Gly  
 385 390 395 400  
 40 TTT GGT GAA AAA ATC GAT GAT ATT GGA GAG TTT AAC TCA GAA AAC TTT 1248  
 Phe Gly Glu Lys Ile Asp Asp Ile Gly Glu Phe Asn Ser Glu Asn Phe  
 405 410 415  
 45 ATG AAA GGT CTC TTG GAA GGT TTA ATC TAA 1278  
 Met Lys Gly Leu Leu Glu Gly Leu Ile \*  
 420 425

50

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:

55

(A) LENGTH: 425 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

-206-

	Met	Gly	Leu	Phe	Asp	Arg	Leu	Phe	Gly	Lys	Lys	Glu	Glu	Pro	Lys	Ile	
	1				5					10					15		
5	Glu	Glu	Val	Val	Lys	Glu	Ala	Leu	Glu	Asn	Leu	Asp	Leu	Ser	Glu	Asp	
				20					25					30			
	Val	Asp	Pro	Thr	Phe	Thr	Glu	Val	Glu	Glu	Val	Ser	Gln	Glu	Glu	Ala	
			35					40					45				
10	Glu	Val	Glu	Ile	Val	Glu	Gln	Ala	Val	Phe	Gln	Glu	Glu	Glu	Ile	Gln	
		50					55					60					
	Asp	Thr	Val	Glu	Glu	Ser	Leu	Asp	Leu	Glu	Pro	Val	Val	Glu	Val	Ser	
15		65				70					75				80		
	Gln	Lys	Glu	Val	Glu	Glu	Phe	Pro	His	Ser	Glu	Glu	Gly	Asn	Thr	Glu	
					85					90					95		
20	Phe	Leu	Glu	Thr	Ile	Glu	Glu	Asn	Asn	Ser	Glu	Val	Leu	Glu	Pro	Glu	
				100				105						110			
	Arg	Pro	Gln	Ala	Glu	Glu	Thr	Val	Gln	Glu	Lys	Tyr	Asp	Arg	Ser	Leu	
			115					120					125				
25	Lys	Lys	Thr	Arg	Thr	Gly	Phe	Gly	Ala	Arg	Leu	Asn	Ala	Phe	Phe	Ala	
		130					135					140					
	Asn	Phe	Arg	Ser	Val	Asp	Glu	Glu	Phe	Phe	Glu	Glu	Leu	Glu	Glu	Leu	
30		145				150					155					160	
	Leu	Ile	Met	Ser	Asp	Val	Gly	Val	Gln	Val	Ala	Ser	Asn	Leu	Thr	Glu	
					165					170					175		
35	Glu	Leu	Arg	Tyr	Glu	Ala	Lys	Leu	Glu	Asn	Ala	Lys	Lys	Pro	Asp	Ala	
				180					185					190			
	Leu	Arg	Arg	Val	Ile	Ile	Glu	Lys	Leu	Val	Glu	Leu	Tyr	Glu	Lys	Asp	
			195					200					205				
40	Gly	Ser	Tyr	Asp	Glu	Ser	Ile	His	Phe	Gln	Asp	Asn	Leu	Thr	Val	Met	
		210					215					220					
	Leu	Phe	Val	Gly	Val	Asn	Gly	Val	Gly	Lys	Thr	Thr	Ser	Ile	Gly	Lys	
45		225				230					235					240	
	Leu	Ala	His	Arg	Tyr	Lys	Gln	Ala	Gly	Lys	Lys	Val	Met	Leu	Val	Ala	
					245					250					255		
50	Ala	Asp	Thr	Phe	Arg	Ala	Gly	Ala	Val	Ala	Gln	Leu	Ala	Glu	Trp	Gly	
				260				265						270			
	Arg	Arg	Val	Asp	Val	Pro	Val	Val	Thr	Gly	Pro	Glu	Lys	Ala	Asp	Pro	
			275					280					285				
55	Ala	Ser	Val	Val	Phe	Asp	Gly	Met	Glu	Arg	Ala	Val	Ala	Glu	Gly	Ile	
		290					295					300					
	Asp	Ile	Leu	Met	Ile	Asp	Thr	Ala	Gly	Arg	Leu	Gln	Asn	Lys	Asp	Asn	
60		305				310					315					320	
	Leu	Met	Ala	Glu	Leu	Glu	Lys	Ile	Gly	Arg	Ile	Ile	Lys	Arg	Val	Val	

-207-

	325		330		335
	Pro Glu Ala	Pro His Glu Thr Phe Leu Ala Leu Asp Ala Ser Thr Gly			
	340	345	350		
5	Gln Asn Ala Leu Val Gln Ala Lys Glu Phe Ser Lys Ile Thr Pro Leu				
	355	360	365		
10	Thr Gly Ile Val Leu Thr Lys Ile Asp Gly Thr Ala Arg Gly Gly Val				
	370	375	380		
	Val Leu Ala Ile Arg Glu Glu Leu Asn Ile Pro Val Lys Leu Ile Gly				
	385	390	395	400	
15	Phe Gly Glu Lys Ile Asp Asp Ile Gly Glu Phe Asn Ser Glu Asn Phe				
	405	410	415		
	Met Lys Gly Leu Leu Glu Gly Leu Ile				
20	420	425			

## (2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 891 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..891

(D) OTHER INFORMATION: H11146

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

	ATG ACA AAG AAA CAA CTT CAC TTG GTG ATT GTG ACA GGG ATG GGT GGC	48
45	Met Thr Lys Lys Gln Leu His Leu Val Ile Val Thr Gly Met Gly Gly	
	1 5 10 15	
	GCA GGG AAA ACT GTA GCC ATT CAG TCC TTC GAG GAT CTA GGT TAT TTC	96
	Ala Gly Lys Thr Val Ala Ile Gln Ser Phe Glu Asp Leu Gly Tyr Phe	
50	20 25 30	
	ACC ATT GAT AAT ATG CCG CCA GCT CTC TTG CCT AAG TTT TTG CAG CTG	144
	Thr Ile Asp Asn Met Pro Pro Ala Leu Leu Pro Lys Phe Leu Gln Leu	
	35 40 45	
55	GTT GAA ATT AAG GAA GAC AAT CCT AAG TTG GCC TTG GTA GTG GAT ATG	192
	Val Glu Ile Lys Glu Asp Asn Pro Lys Leu Ala Leu Val Val Asp Met	
	50 55 60	
60	CGT AGT CGT TCT TTC TTT TCA GAG ATT CAA GCT GTT TTG GAT GAG TTG	240
	Arg Ser Arg Ser Phe Phe Ser Glu Ile Gln Ala Val Leu Asp Glu Leu	
	65 70 75 80	

-208-

	GAA AAT CAA GAT GGT TTG GAT TTC AAA ATC CTC TTT TTG GAT GCG GCT	288
	Glu Asn Gln Asp Gly Leu Asp Phe Lys Ile Leu Phe Leu Asp Ala Ala	
	85 90 95	
5	GAT AAG GAA TTG GTC GCT CGT TAC AAG GAA ACC AGA CGG AGT CAC CCA	336
	Asp Lys Glu Leu Val Ala Arg Tyr Lys Glu Thr Arg Arg Ser His Pro	
	100 105 110	
10	CTA GCA GCA GAC GGT CGT ATT TTA GAT GGA ATC AAG TTG GAA CGT GAA	384
	Leu Ala Ala Asp Gly Arg Ile Leu Asp Gly Ile Lys Leu Glu Arg Glu	
	115 120 125	
15	CTC TTG GCA CCT TTG AAA AAT ATG AGC CAA AAT GTG GTG GAT ACG ACT	432
	Leu Leu Ala Pro Leu Lys Asn Met Ser Gln Asn Val Val Asp Thr Thr	
	130 135 140	
20	GAA CTC ACT CCA CGT GAG CTG CGC AAA ACC CTT GCA GAG CAG TTT TCA	480
	Glu Leu Thr Pro Arg Glu Leu Arg Lys Thr Leu Ala Glu Gln Phe Ser	
	145 150 155 160	
25	GAC CAA GAA CAA GCT CAG TCT TTC CGT ATC GAA GTC ATG TCT TTC GGA	528
	Asp Gln Glu Gln Ala Gln Ser Phe Arg Ile Glu Val Met Ser Phe Gly	
	165 170 175	
30	TTT AAG TAT GGA ATC CCG ATT GAT GCG GAC TTG GTC TTT GAT GTC CGT	576
	Phe Lys Tyr Gly Ile Pro Ile Asp Ala Asp Leu Val Phe Asp Val Arg	
	180 185 190	
35	TTC TTG CCA AAT CCC TAT TAT TTA CCA GAA CTG AGA AAC CAA ACG GGT	624
	Phe Leu Pro Asn Pro Tyr Tyr Leu Pro Glu Leu Arg Asn Gln Thr Gly	
	195 200 205	
40	GTG GAT GAA CCT GTT TAT GAT TAT GTC ATG AAC CAT CCT GAG TCA GAA	672
	Val Asp Glu Pro Val Tyr Asp Tyr Val Met Asn His Pro Glu Ser Glu	
	210 215 220	
45	GAC TTT TAT CAA CAT TTA TTG GCC TTG ATT GAG CCG ATT CTG CCA AGT	720
	Asp Phe Tyr Gln His Leu Leu Ala Leu Ile Glu Pro Ile Leu Pro Ser	
	225 230 235 240	
50	TAC CAA AAG GAA GGT AAG TCC GTT TTG ACC ATT GCC ATG GGA TGT ACG	768
	Tyr Gln Lys Glu Gly Lys Ser Val Leu Thr Ile Ala Met Gly Cys Thr	
	245 250 255	
55	GGT GGA CAA CAC CGT AGT GTG GCA TTT GCT AAA CGC TTG GTG CAG GAC	816
	Gly Gly Gln His Arg Ser Val Ala Phe Ala Lys Arg Leu Val Gln Asp	
	260 265 270	
60	TTA TCC AAG AAT TGG TCT GTT AAT GAA GGG CAT CGC GAC AAA GAC CGC	864
	Leu Ser Lys Asn Trp Ser Val Asn Glu Gly His Arg Asp Lys Asp Arg	
	275 280 285	
65	AGA AAG GAA ACG GTA AAC CGT TCA TGA	891
	Arg Lys Glu Thr Val Asn Arg Ser *	
	290 295	

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:

-209-

(A) LENGTH: 296 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

```

10 Met Thr Lys Lys Gln Leu His Leu Val Ile Val Thr Gly Met Gly Gly
    1          5          10          15

    Ala Gly Lys Thr Val Ala Ile Gln Ser Phe Glu Asp Leu Gly Tyr Phe
    20          25          30

15 Thr Ile Asp Asn Met Pro Pro Ala Leu Leu Pro Lys Phe Leu Gln Leu
    35          40          45

20 Val Glu Ile Lys Glu Asp Asn Pro Lys Leu Ala Leu Val Val Asp Met
    50          55          60

    Arg Ser Arg Ser Phe Phe Ser Glu Ile Gln Ala Val Leu Asp Glu Leu
    65          70          75          80

25 Glu Asn Gln Asp Gly Leu Asp Phe Lys Ile Leu Phe Leu Asp Ala Ala
    85          90          95

    Asp Lys Glu Leu Val Ala Arg Tyr Lys Glu Thr Arg Arg Ser His Pro
    100          105          110

30 Leu Ala Ala Asp Gly Arg Ile Leu Asp Gly Ile Lys Leu Glu Arg Glu
    115          120          125

    Leu Leu Ala Pro Leu Lys Asn Met Ser Gln Asn Val Val Asp Thr Thr
    130          135          140

35 Glu Leu Thr Pro Arg Glu Leu Arg Lys Thr Leu Ala Glu Gln Phe Ser
    145          150          155          160

40 Asp Gln Glu Gln Ala Gln Ser Phe Arg Ile Glu Val Met Ser Phe Gly
    165          170          175

    Phe Lys Tyr Gly Ile Pro Ile Asp Ala Asp Leu Val Phe Asp Val Arg
    180          185          190

45 Phe Leu Pro Asn Pro Tyr Tyr Leu Pro Glu Leu Arg Asn Gln Thr Gly
    195          200          205

    Val Asp Glu Pro Val Tyr Asp Tyr Val Met Asn His Pro Glu Ser Glu
    210          215          220

50 Asp Phe Tyr Gln His Leu Leu Ala Leu Ile Glu Pro Ile Leu Pro Ser
    225          230          235          240

55 Tyr Gln Lys Glu Gly Lys Ser Val Leu Thr Ile Ala Met Gly Cys Thr
    245          250          255

    Gly Gly Gln His Arg Ser Val Ala Phe Ala Lys Arg Leu Val Gln Asp
    260          265          270

60 Leu Ser Lys Asn Trp Ser Val Asn Glu Gly His Arg Asp Lys Asp Arg

```

-210-

275                                      280                                      285

Arg Lys Glu Thr Val Asn Arg Ser  
290                                      295

5

(2) INFORMATION FOR SEQ ID NO:123:

10        (i) SEQUENCE CHARACTERISTICS:  
           (A) LENGTH: 329 amino acids  
           (B) TYPE: amino acid  
           (C) STRANDEDNESS: not relevant  
           (D) TOPOLOGY: not relevant

15        (ii) MOLECULE TYPE: peptide

          (iii) HYPOTHETICAL: NO

20        (iv) ANTI-SENSE: NO

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

25        Met Val Glu Val Pro Asp Glu Arg Leu Gln Lys Leu Thr Glu Met Ile  
           1                                      5                                      10                                      15

          Thr Pro Lys Lys Thr Val Pro Thr Thr Phe Glu Phe Thr Asp Ile Ala  
                                     20                                      25                                      30

30        Gly Ile Val Lys Gly Ala Ser Lys Gly Glu Gly Leu Gly Asn Lys Phe  
                                     35                                      40                                      45

          Leu Ala Asn Ile Arg Glu Val Asp Ala Ile Val His Val Val Arg Ala  
                                     50                                      55                                      60

35        Phe Asp Asp Glu Asn Val Met Arg Glu Gln Gly Arg Glu Asp Ala Phe  
           65                                      70                                      75                                      80

40        Val Asp Pro Leu Ala Asp Ile Asp Thr Ile Asn Leu Glu Leu Ile Leu  
                                     85                                      90                                      95

          Ala Asp Leu Glu Ser Val Asn Lys Arg Tyr Ala Arg Val Glu Lys Met  
                                     100                                      105                                      110

45        Ala Arg Thr Gln Lys Asp Lys Glu Ser Val Ala Glu Phe Asn Val Leu  
                                     115                                      120                                      125

50        Gln Lys Ile Lys Pro Val Leu Glu Asp Gly Lys Ser Ala Arg Thr Ile  
           130                                      135                                      140

          Glu Phe Thr Asp Glu Glu Gln Lys Val Val Lys Gly Leu Phe Leu Leu  
           145                                      150                                      155                                      160

55        Thr Thr Lys Pro Val Leu Tyr Val Ala Asn Val Asp Glu Asp Val Val  
                                     165                                      170                                      175

          Ser Glu Pro Asp Ser Ile Asp Tyr Val Lys Gln Ile Arg Glu Phe Ala  
                                     180                                      185                                      190

60        Ala Thr Glu Asn Ala Glu Val Val Val Ile Ser Ala Arg Ala Glu Glu

-211-

	195	200	205
	Glu Ile Ser Glu Leu Asp Asp	Glu Asp Lys Lys	Glu Phe Leu Glu Ala
	210	215	220
5	Ile Gly Leu Thr Glu Ser Gly Val Asp Lys Leu Thr Arg Ala Ala Tyr		
	225	230	235 240
10	His Leu Leu Gly Leu Gly Thr Tyr Phe Thr Ala Gly Glu Lys Glu Val		
	245	250	255
	Arg Ala Trp Thr Phe Lys Arg Gly Met Lys Ala Pro Gln Ala Ala Gly		
	260	265	270
15	Ile Ile His Ser Asp Phe Glu Lys Gly Phe Ile Arg Ala Val Thr Met		
	275	280	285
	Ser Tyr Glu Asp Leu Val Lys Tyr Gly Ser Glu Lys Ala Val Lys Glu		
	290	295	300
20	Ala Gly Arg Leu Arg Glu Glu Gly Lys Glu Tyr Ile Val Gln Asp Gly		
	305	310	315 320
25	Asp Ile Met Glu Phe Arg Phe Asn Val		
	325		

## (2) INFORMATION FOR SEQ ID NO:124:

30	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 189 amino acids
	(B) TYPE: amino acid
	(C) STRANDEDNESS: not relevant
	(D) TOPOLOGY: not relevant
35	(ii) MOLECULE TYPE: peptide
	(iii) HYPOTHETICAL: NO
40	(iv) ANTI-SENSE: NO
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:
45	Met Ser Ala Ser Glu Gly Arg Asp Pro Tyr Glu Asp Tyr Leu Ala Ile
	1 5 10 15
	Asn Lys Glu Leu Glu Ser Tyr Asn Leu Arg Leu Met Glu Arg Pro Gln
	20 25 30
50	Ile Ile Val Thr Asn Lys Met Asp Met Pro Glu Ser Gln Glu Asn Leu
	35 40 45
	Glu Glu Phe Lys Lys Lys Leu Ala Glu Asn Tyr Asp Glu Phe Glu Glu
	50 55 60
55	Leu Pro Ala Ile Phe Pro Ile Ser Gly Leu Thr Lys Gln Gly Leu Ala
	65 70 75 80
60	Thr Leu Leu Asp Ala Thr Ala Glu Leu Leu Asp Lys Thr Pro Glu Phe
	85 90 95



-212-

Leu Leu Tyr Asp Glu Ser Asp Met Glu Glu Glu Val Tyr Tyr Gly Phe  
 100 105 110  
 5 Asp Glu Glu Glu Lys Ala Phe Glu Ile Ser Arg Asp Asp Ala Thr  
 115 120 125  
 Trp Val Leu Ser Gly Glu Lys Leu Met Lys Leu Phe Asn Met Thr Asn  
 130 135 140  
 10 Phe Asp Arg Asp Glu Ser Val Met Lys Phe Ala Arg Gln Leu Arg Gly  
 145 150 155 160  
 Met Gly Val Asp Glu Ala Leu Arg Ala Arg Gly Ala Lys Asp Gly Asp  
 165 170 175  
 15 Leu Val Arg Ile Gly Lys Phe Glu Phe Glu Phe Val Asp  
 180 185

## (2) INFORMATION FOR SEQ ID NO:125:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 226 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 25 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

35 Met Asn Ile Gln Gln Leu Arg Tyr Val Val Ala Ile Ala Asn Ser Gly  
 1 5 10 15  
 40 Thr Phe Arg Glu Ala Ala Glu Lys Met Tyr Val Ser Gln Pro Ser Leu  
 20 25 30  
 Ser Ile Ser Val Arg Asp Leu Glu Lys Glu Leu Gly Phe Lys Ile Phe  
 35 40 45  
 45 Arg Arg Thr Ser Ser Gly Thr Phe Leu Thr Arg Arg Gly Met Glu Phe  
 50 55 60  
 Tyr Glu Lys Ala Gln Glu Leu Val Lys Gly Phe Asp Ile Phe Gln Asn  
 65 70 75 80  
 50 Gln Tyr Ala Asn Pro Glu Glu Glu Lys Asp Glu Phe Ser Val Ala Ser  
 85 90 95  
 Gln His Tyr Asp Phe Leu Pro Pro Thr Ile Thr Ala Phe Ser Glu Arg  
 100 105 110  
 Tyr Pro Asp Tyr Lys Asn Phe Arg Ile Phe Glu Ser Thr Thr Val Gln  
 115 120 125  
 60 Ile Leu Asp Glu Val Ala Gln Gly His Ser Glu Ile Gly Ile Ile Tyr  
 130 135 140

-213-

Leu Asn Asn Gln Asn Lys Lys Gly Ile Met Gln Arg Val Glu Lys Leu  
 145 150 155 160  
 5 Gly Leu Glu Val Ile Glu Leu Ile Pro Phe His Thr His Ile Tyr Leu  
 165 170 175  
 Cys Glu Gly His Pro Leu Ala Gln Lys Glu Glu Leu Val Met Glu Asp  
 180 185 190  
 10 Leu Ala Asp Leu Pro Thr Val Arg Phe Thr Gln Glu Lys Asp Glu Tyr  
 195 200 205  
 Leu Tyr Tyr Ser Glu Asn Phe Val Asp Thr Ser Ala Thr His Arg Cys  
 210 215 220  
 15 Leu Met  
 225

20 (2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 119 amino acids  
 (B) TYPE: amino acid  
 25 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

30 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

Met Lys Lys Arg Ala Ile Val Ala Val Ile Val Leu Leu Leu Ile Gly  
 1 5 10 15  
 40 Leu Asp Gln Leu Val Lys Ser Tyr Ile Val Gln Gln Ile Pro Leu Gly  
 20 25 30  
 Glu Val Arg Ser Trp Ile Pro Asn Phe Val Ser Leu Thr Tyr Leu Gln  
 35 40 45  
 45 Asn Arg Gly Ala Ala Phe Ser Ile Leu Gln Asp Gln Gln Leu Leu Phe  
 50 55 60  
 50 Ala Val Ile Thr Leu Val Val Val Ile Gly Ala Ile Trp Tyr Leu His  
 65 70 75 80  
 Lys His Met Glu Asp Ser Phe Trp Met Val Leu Gly Leu Thr Leu Ile  
 85 90 95  
 55 Ile Ala Gly Gly Pro Gly Asn Phe Ile Asp Arg Val Ser Gln Gly Phe  
 100 105 110  
 Val Val Asp Met Phe His Leu  
 115  
 60

(2) INFORMATION FOR SEQ ID NO:127:

-214-

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 170 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Lys | Tyr | Leu | Leu | Lys | Leu | Leu | Val | Tyr | Cys | Phe | Ser | Ala | Leu |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| Thr | Phe | Gly | Ser | Leu | Phe | Leu | Ile | Ile | Gly | Phe | Ile | Leu | Ile | Lys | Gly |
|     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |     |
| Leu | Pro | His | Leu | Ser | Leu | Ser | Leu | Phe | Ser | Trp | Thr | Tyr | Thr | Ser | Glu |
|     |     | 35  |     |     |     |     | 40  |     |     |     | 45  |     |     |     |     |
| Asn | Ile | Ser | Leu | Met | Pro | Ala | Ile | Ile | Ser | Thr | Val | Ile | Leu | Val | Phe |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Gly | Ala | Leu | Leu | Leu | Ala | Leu | Pro | Ile | Gly | Ile | Phe | Ala | Gly | Phe | Tyr |
| 65  |     |     |     |     | 70  |     |     |     | 75  |     |     |     |     | 80  |     |
| Leu | Val | Glu | Tyr | Thr | Lys | Lys | Asp | Ser | Leu | Cys | Val | Lys | Ile | Met | Arg |
|     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Leu | Ala | Ser | Asp | Thr | Leu | Ser | Gly | Ile | Pro | Ser | Ile | Val | Phe | Gly | Leu |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
| Phe | Gly | Met | Leu | Phe | Phe | Val | Val | Phe | Leu | Gly | Phe | Gln | Tyr | Ser | Leu |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |
| Leu | Ser | Gly | Ile | Leu | Thr | Ser | Val | Ile | Met | Val | Leu | Pro | Val | Ile | Ile |
|     |     | 130 |     |     |     |     | 135 |     |     |     | 140 |     |     |     |     |
| Arg | Ser | Thr | Glu | Glu | Ala | Leu | Leu | Ser | Val | Ser | Asp | Ser | Met | Arg | Gln |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     | 160 |     |
| Ala | Ser | Tyr | Gly | Leu | Gly | Ala | Leu | Ser | Tyr |     |     |     |     |     |     |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     |     |     |
- (2) INFORMATION FOR SEQ ID NO:128:
- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 154 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO

-215-

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

5 Met Lys Thr Glu Gln Thr Ala Ser Lys Thr Ser Ala Leu Lys Gly Lys  
1 5 10 15

10 Glu Val Ala Asp Phe Glu Leu Met Gly Val Asp Gly Lys Thr Tyr Arg  
20 25 30

Leu Ser Asp Tyr Lys Gly Lys Lys Val Tyr Leu Lys Phe Trp Ala Ser  
35 40 45

15 Trp Cys Ser Ile Cys Leu Ala Ser Leu Pro Asp Thr Asp Glu Ile Ala  
50 55 60

Lys Glu Ala Gly Asp Asp Tyr Val Val Leu Thr Val Val Ser Pro Gly  
65 70 75 80

20 His Lys Gly Glu Gln Ser Glu Ala Asp Phe Lys Asn Trp Tyr Lys Gly  
85 90 95

25 Leu Asp Tyr Lys Asn Leu Pro Val Leu Val Asp Pro Ser Gly Lys Leu  
100 105 110

Leu Glu Thr Tyr Gly Val Arg Ser Tyr Pro Thr Gln Ala Phe Ile Asp  
115 120 125

30 Lys Glu Gly Lys Leu Val Lys Thr His Pro Gly Phe Met Glu Lys Asp  
130 135 140

Ala Ile Leu Gln Thr Leu Lys Glu Leu Ser  
145 150

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 181 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

45

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

55 Met Lys Lys Glu Gln Ile Pro Asn Leu Leu Thr Ile Gly Arg Ile Leu  
1 5 10 15

Phe Ile Pro Ile Phe Ile Phe Ile Leu Thr Ile Gly Asn Ser Ile Glu  
20 25 30

60 Ser His Ile Val Ala Ala Ile Ile Phe Ala Val Ala Ser Ile Thr Asp

-216-

35                      40                      45  
 Tyr Leu Asp Gly Tyr Leu Ala Arg Lys Trp Asn Val Val Ser Asn Phe  
     50                      55                      60  
 5      Gly Lys Phe Ala Asp Pro Met Ala Asp Lys Leu Leu Val Met Ser Ala  
     65                      70                      75                      80  
 10     Phe Ile Met Leu Ile Glu Leu Gly Met Ala Pro Ala Trp Ile Val Ala  
                             85                      90                      95  
 Val Ile Ile Cys Arg Glu Leu Ala Val Thr Gly Leu Arg Leu Leu Leu  
                             100                      105                      110  
 15     Val Glu Thr Gly Gly Thr Ile Leu Ala Ala Ala Met Pro Gly Lys Ile  
                             115                      120                      125  
 Lys Thr Phe Ser Gln Met Phe Ala Ile Ile Phe Leu Leu Leu His Trp  
     130                      135                      140  
 20     Thr Leu Leu Gly Gln Val Leu Leu Tyr Val Ala Leu Phe Phe Thr Ile  
     145                      150                      155                      160  
 25     Tyr Ser Gly Tyr Asp Tyr Phe Lys Gly Ser Ala Tyr Val Phe Lys Gly  
                             165                      170                      175  
 Thr Phe Gly Ser Lys  
                             180

30    (2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 111 amino acids  
     (B) TYPE: amino acid  
 35     (C) STRANDEDNESS: not relevant  
     (D) TOPOLOGY: not relevant

40    (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

50     Leu Arg Leu Lys Glu Met Asn Gly Asp Met Ile His Ala Ala Tyr Asp  
     1                      5                      10                      15  
 Leu Gly Ala Ser Gln Phe Gln Met Phe Lys Glu Ile Met Leu Pro Tyr  
                             20                      25                      30  
 55     Leu Thr Pro Ser Ile Ile Ala Gly Tyr Phe Met Ala Phe Thr Tyr Ser  
                             35                      40                      45  
 Leu Asp Asp Phe Ala Val Thr Phe Phe Val Thr Gly Asn Gly Phe Ser  
     50                      55                      60  
 60     Thr Leu Ser Val Glu Ile Tyr Ser Arg Ala Arg Lys Gly Ile Ser Leu

-217-

65                                      70                                      75                                      80  
 Glu Ile Asn Ala Leu Ser Ala Leu Val Phe Leu Phe Ser Ile Ile Leu  
    85                                      90                                      95  
 5 Val Val Gly Tyr Tyr Phe Ile Ser Arg Glu Lys Glu Glu Gln Ala  
    100                                      105                                      110

## (2) INFORMATION FOR SEQ ID NO:131:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

Pro Gln Phe Thr Glu Glu Thr Gly Ile Gln Val Gln Tyr Glu Ala Phe  
 1                                      5                                      10                                      15  
 Asp Ser Asn Glu Ala Met Tyr Thr Lys Ile Lys Gln Gly Gly Thr Thr  
 20                                      25                                      30  
 Tyr Asp Ile Ala Ile Pro Ser Glu Tyr Met Ile Asn Lys Met Lys Asp  
 35                                      40                                      45  
 Glu Asp Leu Leu Val Pro Leu Asp Tyr Ser Lys  
 50                                      55

## (2) INFORMATION FOR SEQ ID NO:132:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 232 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Met Gln Thr Gln Glu Lys His Ser Gln Ala Ala Val Leu Gly Leu Gln  
 1                                      5                                      10                                      15  
 His Leu Leu Ala Met Tyr Ser Gly Ser Ile Leu Val Pro Ile Met Ile  
 20                                      25                                      30  
 Ala Thr Ala Leu Gly Tyr Ser Ala Glu Gln Leu Thr Tyr Leu Ile Ser

-218-

	35	40	45
5	Thr Asp Ile Phe Met Cys Gly Val Ala Thr Phe Leu Gln Leu Gln Leu 50 55 60		
	Asn Lys Tyr Phe Gly Ile Gly Leu Pro Val Val Leu Gly Val Ala Phe 65 70 75 80		
10	Gln Ser Val Ala Pro Leu Ile Met Ile Gly Gln Ser His Gly Ser Gly 85 90 95		
	Ala Met Phe Gly Ala Leu Ile Ala Ser Gly Ile Tyr Val Val Leu Val 100 105 110		
15	Ser Gly Ile Phe Ser Lys Val Ala Asn Leu Phe Pro Ser Ile Val Thr 115 120 125		
	Gly Ser Val Ile Thr Thr Ile Gly Leu Thr Leu Ile Pro Val Ala Ile 130 135 140		
20	Gly Asn Met Gly Asn Asn Val Pro Glu Pro Thr Gly Gln Ser Leu Leu 145 150 155 160		
	Leu Ala Ala Ile Thr Val Leu Ile Ile Leu Leu Ile Asn Ile Phe Thr 165 170 175		
25	Lys Gly Phe Ile Lys Ser Ile Ser Ile Leu Ile Gly Leu Val Val Gly 180 185 190		
	Thr Ala Ile Ala Ala Thr Met Gly Leu Val Asp Phe Ser Pro Val Ala 195 200 205		
30	Val Val His Leu Ser Met Ser Gln Leu His Ser Thr Leu Gly Cys Gln 210 215 220		
35	Pro Leu Lys Ser His Leu Leu Ser 225 230		

## (2) INFORMATION FOR SEQ ID NO:133:

40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 343 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

55	Lys Val Pro Val Tyr Leu Gly Ser Ser Phe Ala Phe Ile Thr Ala Met 1 5 10 15
60	Ser Leu Ala Met Lys Glu Met Gly Gly Asp Val Ser Ala Ala Gln Thr 20 25 30

-219-

Gly Val Ile Leu Thr Gly Leu Val Tyr Val Leu Val Ala Thr Ser Ile  
 35 40 45  
 5 Arg Phe Val Gly Thr Lys Trp Ile Asp Lys Leu Leu Pro Pro Ile Ile  
 50 55 60  
 Ile Gly Pro Met Ile Ile Val Ile Gly Leu Gly Leu Ala Gly Ser Ala  
 65 70 75 80  
 10 Val Thr Asn Ala Gly Leu Val Ala Asp Gly Asn Trp Lys Asn Ala Leu  
 85 90 95  
 Val Ala Val Val Thr Phe Leu Ile Ala Ala Phe Ile Asn Thr Lys Gly  
 100 105 110  
 15 Lys Gly Phe Leu Arg Ile Ile Pro Phe Leu Phe Ala Ile Ile Gly Gly  
 115 120 125  
 Tyr Leu Phe Ala Leu Thr Leu Gly Leu Val Asp Phe Thr Pro Val Leu  
 130 135 140  
 20 Lys Ala Asn Trp Phe Glu Ile Pro Gly Phe Tyr Leu Pro Phe Ser Thr  
 145 150 155 160  
 25 Gly Gly Ala Phe Lys Glu Tyr Asn Leu Tyr Phe Gly Pro Glu Ala Ile  
 165 170 175  
 Ala Ile Leu Pro Ile Ala Ile Val Thr Ile Ser Glu His Ile Gly Asp  
 180 185 190  
 30 His Thr Val Leu Gly Gln Ile Cys Gly Arg Gln Phe Leu Lys Glu Pro  
 195 200 205  
 Gly Leu His Arg Thr Leu Leu Gly Asp Gly Ile Ala Thr Ser Val Ser  
 210 215 220  
 35 Ala Phe Leu Gly Gly Pro Ala Asn Thr Thr Tyr Gly Glu Asn Thr Gly  
 225 230 235 240  
 40 Val Ile Gly Met Thr Arg Ile Ala Ser Val Ser Val Ile Arg Asn Ala  
 245 250 255  
 Ala Phe Ile Ala Ile Ala Leu Ser Phe Leu Gly Lys Phe Thr Ala Leu  
 260 265 270  
 45 Ile Ser Thr Ile Pro Asn Ala Val Leu Gly Gly Met Ser Ile Leu Leu  
 275 280 285  
 Tyr Gly Val Ile Ala Ser Asn Gly Leu Lys Val Leu Ile Lys Glu Arg  
 290 295 300  
 50 Val Asp Phe Ala Gln Met Arg Asn Leu Ile Ile Ala Ser Ala Met Leu  
 305 310 315 320  
 55 Val Leu Gly Leu Gly Gly Ala Ile Leu Lys Leu Gly Pro Val His Phe  
 325 330 335  
 Gln Val Leu Pro Phe Gln Pro  
 340  
 60

(2) INFORMATION FOR SEQ ID NO:134:



-220-

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 184 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Leu | Ile | Ile | Ala | Leu | Ala | Thr | Thr | Leu | Ile | Ala | Ile | Ile | Ile | Ser |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| Ala | Met | Ala | Ala | Tyr | Gly | Ile | Val | Arg | Phe | Phe | Pro | Lys | Leu | Gly | Ala |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
| Ile | Met | Ser | Arg | Leu | Leu | Val | Ile | Thr | Tyr | Ile | Phe | Pro | Pro | Ile | Leu |
|     |     | 35  |     |     |     | 40  |     |     |     |     | 45  |     |     |     |     |
| Leu | Ala | Ile | Pro | Tyr | Ser | Ile | Ala | Ile | Ala | Lys | Val | Gly | Leu | Thr | Asn |
|     | 50  |     |     |     | 55  |     |     |     |     |     | 60  |     |     |     |     |
| Ser | Leu | Phe | Gly | Leu | Met | Met | Val | Tyr | Leu | Ser | Phe | Ser | Val | Pro | Tyr |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Ala | Val | Trp | Leu | Leu | Val | Gly | Phe | Phe | Gln | Thr | Val | Pro | Ile | Gly | Ile |
|     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Glu | Glu | Ala | Ala | Arg | Ile | Asp | Gly | Ala | Asn | Lys | Phe | Val | Thr | Phe | Tyr |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
| Lys | Val | Val | Leu | Pro | Ile | Val | Ala | Pro | Gly | Ile | Val | Ala | Thr | Ala | Ile |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |
| Tyr | Thr | Phe | Ile | Asn | Ala | Trp | Asn | Glu | Phe | Leu | Tyr | Ala | Leu | Ile | Leu |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |
| Ile | Asn | Asn | Thr | Gly | Lys | Met | Thr | Val | Ala | Val | Ala | Leu | Arg | Ser | Leu |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |
| Asn | Gly | Ser | Glu | Ile | Leu | Asp | Trp | Gly | Asp | Met | Met | Ala | Ala | Ser | Val |
|     |     |     | 165 |     |     |     |     | 170 |     |     |     |     |     | 175 |     |
| Ile | Val | Val | Leu | Pro | Ser | Ile | Ile |     |     |     |     |     |     |     |     |
|     |     |     | 180 |     |     |     |     |     |     |     |     |     |     |     |     |
- (2) INFORMATION FOR SEQ ID NO:135:
- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 155 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant
- (ii) MOLECULE TYPE: peptide

-221-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Asp Glu Leu Ala Asp Leu Met Met Val Ala Ser Lys Glu Val Glu Asp  
 1 5 10 15  
 Ala Ile Ile Arg Leu Gly Gln Lys Ala Arg Ala Ala Gly Ile His Met  
 20 25 30  
 Ile Leu Ala Thr Gln Arg Pro Ser Val Asp Val Ile Ser Gly Leu Ile  
 35 40 45  
 Lys Ala Asn Val Pro Ser Arg Val Ala Phe Ala Val Ser Ser Gly Thr  
 50 55 60  
 Asp Ser Arg Thr Ile Leu Asp Glu Asn Gly Ala Glu Lys Leu Leu Gly  
 65 70 75 80  
 Arg Gly Asp Met Leu Phe Lys Pro Ile Asp Glu Asn His Pro Val Arg  
 85 90 95  
 Leu Gln Gly Ser Phe Ile Ser Asp Asp Asp Val Glu Arg Ile Val Asn  
 100 105 110  
 Phe Ile Lys Thr Gln Ala Asp Ala Asp Tyr Asp Glu Ser Phe Asp Pro  
 115 120 125  
 Gly Glu Val Ser Glu Asn Glu Gly Glu Phe Ser Asp Gly Asp Ala Gly  
 130 135 140  
 Gly Asp Pro Leu Phe Glu Glu Ala Lys Ser Leu  
 145 150 155

(2) INFORMATION FOR SEQ ID NO:136:

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

45

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

50

(iv) ANTI-SENSE: NO

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

Met Thr Glu Asn Thr Pro Lys Ala Leu Val Gln Val Asn Gln Lys Pro  
 1 5 10 15  
 Leu Ile Glu Tyr Gln Ile Glu Phe Leu Lys Glu Lys Gly Ile Asn Asp  
 20 25 30

-222-

Ile Ile Ile Ile Val Gly Tyr Leu Lys Glu Gln Phe Asp Tyr Leu Lys  
           35                  40                  45  
 5 Glu Lys Tyr Gly Val Arg Leu Val Phe Asn Asp Lys Tyr Ala Asp Tyr  
       50                  55                  60  
 Asn Asn Phe Tyr Ser Leu Tyr Leu Val Lys Glu Glu Leu Ala Asn Ser  
 65                  70                  75                  80  
 10 Tyr Val Ile Asp Ala Asp Asn Tyr Leu Phe Lys Asn Met Phe Arg Asn  
           85                  90                  95  
 Asp Leu Thr Arg Ser Thr Tyr Phe Ser Val Tyr Arg Glu Asp Cys Thr  
 15                  100                  105                  110  
 Asn Glu Trp Phe Leu Val Tyr Gly Asp Asp Tyr Lys Val Gln Asp Ile  
           115                  120                  125  
 20 Ile Val Asp Ser Lys Ala Gly Arg Ile Leu Ser Gly Val Ser Phe Trp  
       130                  135                  140  
 Asp Ala Pro Thr Ala Glu Lys Ile Val Ser  
 25 145                  150

## (2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 286 amino acids  
 30 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

35 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

45 Met Ser Asp Asn Ser Lys Thr Arg Val Val Val Gly Met Ser Gly Gly  
       1                  5                  10                  15  
 Val Asp Ser Ser Val Thr Ala Leu Leu Lys Glu Gln Gly Tyr Asp  
 50                  20                  25                  30  
 Val Ile Gly Ile Phe Met Lys Asn Trp Asp Asp Thr Asp Glu Asn Gly  
       35                  40                  45  
 Val Cys Thr Ala Thr Glu Asp Tyr Lys Asp Val Val Ala Val Ala Asp  
 55                  50                  55                  60  
 Gln Ile Gly Ile Pro Tyr Tyr Ser Val Asn Phe Glu Lys Glu Tyr Trp  
       65                  70                  75                  80  
 60 Asp Arg Val Phe Glu Tyr Phe Leu Ala Glu Tyr Arg Ala Gly Arg Thr  
           85                  90                  95

-223-

Pro Asn Pro Asp Val Met Cys Asn Lys Glu Ile Lys Phe Lys Ala Phe  
                   100                  105                  110  
 5 Leu Asp Tyr Ala Met Thr Leu Gly Ala Asp Tyr Val Ala Thr Gly His  
                   115                  120                  125  
 Tyr Ala Arg Val Ala Arg Asp Glu Asp Gly Thr Val His Met Leu Arg  
           130                  135                  140  
 10 Gly Val Asp Asn Gly Lys Asp Gln Thr Tyr Phe Leu Ser Gln Leu Ser  
       145                  150                  155                  160  
 Gln Glu Gln Leu Gln Lys Thr Met Phe Pro Leu Gly His Leu Lys Lys  
                   165                  170                  175  
 Pro Glu Val Arg Lys Leu Ala Glu Glu Ala Gly Leu Ser Thr Ala Lys  
                   180                  185                  190  
 20 Lys Lys Asp Ser Thr Gly Ile Cys Phe Ile Gly Glu Lys Asn Phe Lys  
           195                  200                  205  
 Asn Phe Leu Ser Asn Tyr Leu Pro Ala Gln Pro Gly Arg Met Met Thr  
       210                  215                  220  
 25 Val Asp Gly Arg Asp Met Gly Glu His Ala Gly Leu Met Tyr Tyr Thr  
       225                  230                  235                  240  
 Ile Gly Gln Arg Gly Gly Leu Gly Ile Gly Gly Gln His Gly Gly Asp  
                   245                  250                  255  
 Asn Ala Pro Trp Phe Val Val Gly Lys Asp Leu Ser Lys Asn Ile Leu  
                   260                  265                  270  
 35 Tyr Val Gly Gln Gly Phe Tyr His Asp Ser Leu Met Ser Thr  
           275                  280                  285

## (2) INFORMATION FOR SEQ ID NO:138:

40 (i) SEQUENCE CHARACTERISTICS:  
       (A) LENGTH: 648 amino acids  
       (B) TYPE: amino acid  
       (C) STRANDEDNESS: not relevant  
       (D) TOPOLOGY: not relevant

45 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

55 Met Glu Val Phe Glu Ser Leu Lys Ala Asn Leu Val Gly Lys Asn Ala  
       1                  5                  10                  15  
 Arg Ile Val Leu Pro Glu Gly Glu Glu Pro Arg Ile Leu Gln Ala Thr  
           20                  25                  30  
 60 Lys Arg Leu Val Lys Glu Thr Glu Val Ile Pro Val Leu Leu Gly Asn

-224-

	35	40	45
	Pro Glu Lys Ile Lys Ile Tyr Leu Glu Ile Glu Gly Ile Met Asp Gly		
5	50	55	60
	Tyr Glu Val Ile Asp Pro Gln His Tyr Pro Gln Phe Glu Glu Met Val		
	65	70	75 80
10	Ser Ala Leu Val Glu Arg Arg Lys Gly Lys Met Thr Glu Glu Asp Val		
	85	90	95
	Arg Lys Val Leu Val Glu Asp Val Asn Tyr Phe Gly Val Met Leu Val		
	100	105	110
15	Tyr Leu Gly Leu Val Asp Gly Met Val Ser Gly Ala Ile His Ser Thr		
	115	120	125
	Ala Ser Thr Val Arg Pro Ala Leu Gln Ile Ile Lys Thr Arg Pro Asn		
20	130	135	140
	Val Thr Arg Thr Ser Gly Ala Phe Leu Met Val Arg Gly Thr Glu Arg		
	145	150	155 160
25	Tyr Leu Phe Gly Asp Cys Ala Ile Asn Ile Asn Pro Asp Ala Glu Ala		
	165	170	175
	Leu Ala Glu Ile Ala Ile Asn Ser Ala Ile Thr Ala Lys Met Phe Gly		
	180	185	190
30	Ile Glu Pro Lys Ile Ala Met Leu Ser Tyr Ser Thr Lys Gly Ser Gly		
	195	200	205
	Phe Gly Glu Ser Val Asp Lys Val Val Glu Ala Thr Lys Ile Ala His		
35	210	215	220
	Asp Leu Arg Pro Asp Leu Glu Ile Asp Gly Glu Leu Gln Phe Asp Ala		
	225	230	235 240
40	Ala Phe Val Pro Glu Thr Ala Ala Leu Lys Ala Pro Gly Ser Thr Val		
	245	250	255
	Ala Gly Gln Ala Asn Val Phe Ile Phe Pro Gly Ile Glu Ala Gly Asn		
	260	265	270
45	Ile Gly Tyr Lys Met Ala Glu Arg Leu Gly Gly Phe Ala Ala Val Gly		
	275	280	285
	Pro Val Leu Gln Gly Leu Asn Lys Pro Val Asn Asp Leu Ser Arg Gly		
50	290	295	300
	Cys Asn Ala Asp Asp Val Tyr Lys Leu Thr Leu Ile Thr Ala Ala Gln		
	305	310	315 320
55	Ala Val His Gln Met Glu Val Phe Glu Ser Leu Lys Ala Asn Leu Val		
	325	330	335
	Gly Lys Asn Ala Arg Ile Val Leu Pro Glu Gly Glu Glu Pro Arg Ile		
	340	345	350
60	Leu Gln Ala Thr Lys Arg Leu Val Lys Glu Thr Glu Val Ile Pro Val		
	355	360	365

-225-

Leu Leu Gly Asn Pro Glu Lys Ile Lys Ile Tyr Leu Glu Ile Glu Gly  
 370 375 380  
 5 Ile Met Asp Gly Tyr Glu Val Ile Asp Pro Gln His Tyr Pro Gln Phe  
 385 390 395 400  
 Glu Glu Met Val Ser Ala Leu Val Glu Arg Arg Lys Gly Lys Met Thr  
 405 410 415  
 10 Glu Glu Asp Val Arg Lys Val Leu Val Glu Asp Val Asn Tyr Phe Gly  
 420 425 430  
 Val Met Leu Val Tyr Leu Gly Leu Val Asp Gly Met Val Ser Gly Ala  
 435 440 445  
 15 Ile His Ser Thr Ala Ser Thr Val Arg Pro Ala Leu Gln Ile Ile Lys  
 450 455 460  
 Thr Arg Pro Asn Val Thr Arg Thr Ser Gly Ala Phe Leu Met Val Arg  
 465 470 475 480  
 Gly Thr Glu Arg Tyr Leu Phe Gly Asp Cys Ala Ile Asn Ile Asn Pro  
 485 490 495  
 25 Asp Ala Glu Ala Leu Ala Glu Ile Ala Ile Asn Ser Ala Ile Thr Ala  
 500 505 510  
 Lys Met Phe Gly Ile Glu Pro Lys Ile Ala Met Leu Ser Tyr Ser Thr  
 515 520 525  
 Lys Gly Ser Gly Phe Gly Glu Ser Val Asp Lys Val Val Glu Ala Thr  
 530 535 540  
 35 Lys Ile Ala His Asp Leu Arg Pro Asp Leu Glu Ile Asp Gly Glu Leu  
 545 550 555 560  
 Gln Phe Asp Ala Ala Phe Val Pro Glu Thr Ala Ala Leu Lys Ala Pro  
 565 570 575  
 40 Gly Ser Thr Val Ala Gly Gln Ala Asn Val Phe Ile Phe Pro Gly Ile  
 580 585 590  
 Glu Ala Gly Asn Ile Gly Tyr Lys Met Ala Glu Arg Leu Gly Gly Phe  
 595 600 605  
 45 Ala Ala Val Gly Pro Val Leu Gln Gly Leu Asn Lys Pro Val Asn Asp  
 610 615 620  
 Leu Ser Arg Gly Cys Asn Ala Asp Asp Val Tyr Lys Leu Thr Leu Ile  
 625 630 635 640  
 Thr Ala Ala Gln Ala Val His Gln  
 645  
 55

(2) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 135 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

60

-226-

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

Met Arg Asn Leu Lys Ser Ile Leu Arg Arg His Ile Ser Leu Leu Gly  
 1 5 10 15

15 Phe Leu Gly Val Leu Ser Ile Trp Gln Leu Ala Gly Phe Leu Lys Leu  
 20 25 30

Leu Pro Lys Phe Ile Leu Pro Thr Pro Leu Glu Ile Leu Gln Pro Phe  
 35 40 45

20 Val Arg Asp Arg Glu Phe Leu Trp His His Ser Trp Ala Thr Leu Arg  
 50 55 60

25 Val Ala Leu Leu Gly Leu Ile Leu Gly Val Leu Ile Ala Cys Leu Met  
 65 70 75 80

Ala Val Leu Met Asp Ser Leu Thr Trp Leu Asn Asp Leu Ile Tyr Pro  
 85 90 95

30 Met Met Val Val Ile Gln Thr Ile Pro Thr Ile Ala Ile Ala Pro Ile  
 100 105 110

Leu Val Leu Trp Leu Gly Tyr Gly Ile Phe Ala Gln Asp Cys Leu Asp  
 115 120 125

35 Tyr Leu Asn Asn Asn Leu Ser  
 130 135

(2) INFORMATION FOR SEQ ID NO:140:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 279 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

45 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

50

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

55

Pro Trp Ser Leu Val Asp Glu Tyr Glu Gln Leu Tyr Ala Thr Ile Gly  
 1 5 10 15

60 Trp His Pro Thr Glu Ala Gly Thr Tyr Thr Glu Glu Val Glu Ala Tyr  
 20 25 30

-227-

Leu Leu Asp Lys Leu Lys His Ser Lys Val Val Ala Leu Gly Glu Ile  
           35                          40                          45  
 Gly Leu Asp Tyr His Trp Met Thr Ala Pro Glu Val Gln Glu Gln Val  
 5           50                          55                          60  
 Phe Arg Arg Gln Ile Gln Leu Ser Lys Asp Leu Asp Leu Pro Phe Val  
           65                          70                          75                          80  
 Val His Thr Arg Asp Ala Leu Glu Asp Thr Tyr Glu Ile Ile Lys Ser  
 10                           85                          90                          95  
 Glu Gly Val Gly Pro Arg Gly Gly Ile Met His Ser Phe Ser Gly Thr  
                           100                          105                          110  
 Leu Glu Trp Ala Arg Tyr Arg Asp Leu Gly Met Thr Ile Ser Phe Ser  
                           115                          120                          125  
 Gly Val Val Thr Phe Lys Lys Ala Thr Asp Leu Gln Glu Ala Ala Lys  
 20           130                          135                          140  
 Glu Leu Pro Leu Asp Lys Met Leu Val Glu Thr Asp Ala Pro Tyr Leu  
           145                          150                          155                          160  
 Ala Pro Val Pro Lys Arg Gly Arg Glu Asn Lys Thr Ala Tyr Thr Arg  
 25                           165                          170                          175  
 Tyr Val Val Asp Phe Ile Ala Asp Leu Arg Gly Met Thr Thr Glu Glu  
                           180                          185                          190  
 Leu Ala Val Ala Thr Thr Ala Asn Ala Glu Arg Ile Phe Gly Ile Gly  
                           195                          200                          205  
 Gln Gln Val Met Lys Glu Arg Ile Ser Gln Val Ile Val Val Glu Gly  
 35           210                          215                          220  
 Arg Asp Asp Thr Val Asn Leu Lys Arg Tyr Phe Asp Val Glu Thr Tyr  
           225                          230                          235                          240  
 Glu Thr Arg Gly Ser Ala Ile Asn Asp Gln Asp Ile Glu Arg Ile Gln  
 40                           245                          250                          255  
 Arg Leu His Gln Arg His Gly Val Ile Val Phe Thr Asp Pro Asp Phe  
                           260                          265                          270  
 Asn Gly Asp Gly Phe Gly Ala  
           275

## (2) INFORMATION FOR SEQ ID NO:141:

- 50 (i) SEQUENCE CHARACTERISTICS:  
       (A) LENGTH: 147 amino acids  
       (B) TYPE: amino acid  
       (C) STRANDEDNESS: not relevant  
 55 (D) TOPOLOGY: not relevant  
  
       (ii) MOLECULE TYPE: peptide  
  
       (iii) HYPOTHETICAL: NO  
 60 (iv) ANTI-SENSE: NO



-228-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

5 Met Lys Ile Ile Ile Gln Arg Val Lys Lys Ala Gln Val Ser Ile Glu  
 1 5 10 15  
 Gly Gln Ile Gln Gly Lys Ile Asn Gln Gly Leu Leu Leu Leu Val Gly  
 20 25 30  
 10 Val Gly Pro Glu Asp Gln Glu Glu Asp Leu Asp Tyr Ala Val Arg Lys  
 35 40 45  
 15 Leu Val Asn Met Arg Ile Phe Ser Asp Ala Glu Gly Lys Met Asn Leu  
 50 55 60  
 Ser Val Lys Asp Ile Glu Gly Glu Ile Leu Ser Ile Ser Gln Phe Thr  
 65 70 75 80  
 20 Leu Phe Ala Asp Thr Lys Lys Gly Asn Arg Pro Ala Phe Thr Gly Ala  
 85 90 95  
 Ala Lys Pro Asp Met Ala Ser Asp Phe Tyr Asp Ala Phe Asn Gln Lys  
 100 105 110  
 25 Leu Ala Gln Glu Val Pro Val Gln Thr Gly Ile Phe Gly Ala Asp Met  
 115 120 125  
 30 Gln Val Glu Leu Val Asn Asn Gly Pro Val Thr Ile Ile Leu Asp Thr  
 130 135 140  
 Lys Lys Arg  
 145

## 35 (2) INFORMATION FOR SEQ ID NO:142:

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 238 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

45 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

Met Ile Leu Ser Met Val Ser Thr Pro Leu Pro Ser Ser Pro Cys Lys  
 1 5 10 15  
 55 Tyr Arg Lys Gln Leu Tyr Leu Gln Glu Asp Leu Arg Gly Lys Asn Val  
 20 25 30  
 Glu Lys Val Lys Glu Leu Ala Thr Glu Lys Lys Val Ser Ile Ser Trp  
 35 40 45  
 60 Thr Ser Lys Lys Ser Leu Ser Glu Met Thr Glu Gly Ala Val His Gln

-229-

	50		55		60
	Gly Phe,Val Leu Arg Val Ser Glu Phe Ala Tyr Ser Glu Leu Asp Tyr				
	65		70		75 80
5	Ile Leu Ala Lys Thr Arg Gln Glu Glu Asn Pro Leu Leu Leu Ile Leu				
		85		90	95
10	Asp Gly Leu Thr Asp Pro His Asn Leu Gly Ser Ile Leu Arg Thr Ala				
		100		105	110
	Asp Ala Thr Asn Val Ser Gly Val Ile Ile Pro Lys His Arg Ala Val				
		115		120	125
15	Gly Val Thr Pro Val Val Ala Lys Thr Ala Thr Gly Ala Ile Glu His				
		130		135	140
	Val Pro Ile Ala Arg Val Thr Asn Leu Ser Gln Thr Leu Asp Lys Leu				
		145		150	155 160
20	Lys Asp Glu Gly Phe Trp Thr Phe Gly Thr Asp Met Asn Gly Thr Pro				
		165		170	175
25	Cys His Lys Trp Asn Thr Lys Gly Lys Ile Ala Leu Ile Ile Gly Asn				
		180		185	190
	Glu Gly Lys Gly Ile Ser Ser Asn Ile Lys Lys Gln Val Asp Glu Met				
		195		200	205
30	Ile Thr Ile Pro Met Asn Gly His Val Gln Ser Leu Asn Ala Ser Val				
		210		215	220
	Ala Ala Ala Ile Leu Met Tyr Glu Val Phe Arg Asn Arg Leu				
		225		230	235

## (2) INFORMATION FOR SEQ ID NO:143:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 345 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

	Met Val Gln Gln Ala Ala Thr Val Ser Leu Met Val Leu Phe Leu Val		
	1	5	10 15
55	Pro Gln Leu Arg Asn Ala Tyr Gly Thr Ala Ala Ile Gly Ile Ile Cys		
		20	25 30
60	Gly Leu Tyr Trp Ala Val Ser Ser Asn Met Thr Val Glu Ala Thr Gln		
		35	40 45

-230-

Arg Leu Thr Gly Gly Gly Gly Phe Ala Ile Gly His Gln Gln Gln Phe  
 50 55 60  
 5 Ala Ile Trp Phe Val Asp Lys Val Ala Gly Arg Phe Gly Lys Lys Glu  
 65 70 75 80  
 Glu Ser Leu Asp Asn Leu Lys Leu Pro Lys Phe Leu Ser Ile Phe His  
 85 90 95  
 10 Asp Thr Val Val Ala Ser Ala Thr Leu Met Leu Val Phe Phe Gly Ala  
 100 105 110  
 Ile Leu Leu Ile Leu Gly Pro Asp Ile Met Ser Asn Lys Glu Val Ile  
 115 120 125  
 15 Thr Ser Gly Thr Leu Phe Asn Pro Ala Lys Gln Asp Phe Phe Met Tyr  
 130 135 140  
 20 Ile Ile Gln Thr Ala Phe Thr Phe Ser Val Tyr Leu Phe Val Leu Met  
 145 150 155 160  
 Gln Gly Val Arg Met Phe Val Ser Glu Leu Thr Asn Ala Phe Gln Gly  
 165 170 175  
 25 Ile Ser Asn Lys Leu Leu Pro Gly Ser Phe Pro Ala Val Asp Val Ala  
 180 185 190  
 Ala Ser Tyr Gly Phe Gly Ser Pro Asn Ala Val Leu Ser Gly Phe Thr  
 195 200 205  
 30 Phe Gly Leu Ile Gly Gln Leu Ile Thr Ile Val Leu Leu Ile Val Phe  
 210 215 220  
 Lys Asn Pro Ile Leu Ile Ile Thr Gly Phe Val Pro Val Phe Phe Asp  
 225 230 235 240  
 Asn Ala Ala Ile Ala Val Tyr Ala Asp Lys Arg Gly Gly Trp Lys Ala  
 245 250 255  
 40 Ala Val Ile Leu Ser Phe Ile Ser Gly Val Leu Gln Val Ala Leu Gly  
 260 265 270  
 Ala Leu Cys Val Ala Leu Leu Asp Leu Ala Ser Tyr Gly Gly Tyr His  
 275 280 285  
 45 Gly Asn Ile Asp Phe Glu Phe Pro Trp Leu Gly Phe Gly Tyr Ile Phe  
 290 295 300  
 Lys Tyr Leu Gly Ile Val Gly Tyr Val Leu Val Cys Leu Phe Leu Leu  
 305 310 315 320  
 Val Ile Pro Gln Leu Gln Phe Ala Lys Ala Lys Asp Lys Glu Lys Tyr  
 325 330 335  
 55 Tyr Asn Gly Glu Val Gln Glu Glu Ala  
 340 345

(2) INFORMATION FOR SEQ ID NO:144:

60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 287 amino acids

-231-

(B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

5 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 10 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

15 Met Val Arg Pro Ile Gly Ile Tyr Glu Lys Ala Thr Pro Thr His Phe  
 1 5 10 15

20 Thr Trp Leu Glu Arg Leu Asn Phe Ala Lys Glu Leu Gly Phe Asp Phe  
 20 25 30

Val Glu Met Ser Ile Asp Glu Arg Asp Glu Arg Leu Ala Arg Leu Asp  
 35 40 45

25 Trp Ser Lys Glu Glu Arg Leu Glu Val Val Lys Ala Ile Tyr Glu Thr  
 50 55 60

Gly Val Arg Ile Pro Ser Ile Cys Phe Ser Gly His Arg Arg Tyr Pro  
 65 70 75 80

30 Leu Gly Ser Lys Asp Pro Val Leu Glu Glu Lys Ser Leu Glu Leu Met  
 85 90 95

Lys Lys Cys Ile Glu Leu Ala Gln Asp Leu Gly Val Arg Thr Ile Gln  
 100 105 110

35 Leu Ala Gly Tyr Asp Val Tyr Tyr Glu Glu Lys Ser Pro Gln Thr Arg  
 115 120 125

40 Gln Arg Phe Ile Lys Asn Leu Arg Lys Ala Cys Asp Trp Ala Glu Glu  
 130 135 140

Ala Gln Val Val Leu Ala Ile Glu Ile Met Asp Asp Pro Phe Ile Asn  
 145 150 155 160

45 Ser Ile Glu Lys Tyr Leu Ala Ile Glu Lys Glu Ile Asp Ser Pro Phe  
 165 170 175

50 Leu Phe Val Tyr Pro Asp Ile Gly Asn Val Ser Ala Trp His Asn Asp  
 180 185 190

Ile Tyr Ser Glu Phe Tyr Leu Gly His His Ala Ile Ala Ala Leu His  
 195 200 205

55 Leu Lys Asp Thr Tyr Ala Val Thr Glu Ser Ser Lys Gly Gln Phe Arg  
 210 215 220

Asp Val Pro Phe Gly Gln Gly Cys Val Lys Trp Glu Glu Ala Phe Asp  
 225 230 235 240

60 Ile Leu Lys Glu Thr Asn Tyr Asn Gly Pro Phe Leu Ile Glu Met Trp

-232-

```

245                      250                      255
Ser Glu Asn Cys Glu Thr Val Glu Glu Thr Arg Ala Ala Val Gln Glu
      260                      265                      270
5
Ala Gln Ala Phe Leu Tyr Pro Leu Ile Lys Lys Ala Gly Leu Met
      275                      280                      285

(2) INFORMATION FOR SEQ ID NO:145:

10 (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 221 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: not relevant
15      (D) TOPOLOGY: not relevant

      (ii) MOLECULE TYPE: peptide

20      (iii) HYPOTHETICAL: NO

      (iv) ANTI-SENSE: NO

25      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

Met Thr Lys Arg Ile Pro Asn Leu Gln Val Ala Leu Asp His Ser Asp
1      5      10      15
Leu Gln Gly Ala Ile Lys Ala Ala Val Ser Val Gly Gln Glu Val Asp
30      20      25      30
Ile Ile Glu Ala Gly Thr Val Cys Leu Leu Gln Val Gly Ser Glu Leu
      35      40      45
Ala Glu Val Leu Arg Ser Leu Phe Pro Asp Lys Ile Ile Val Ala Asp
35      50      55      60
Thr Lys Cys Ala Asp Ala Gly Gly Thr Val Ala Lys Asn Asn Ala Val
40      65      70      75      80
Arg Gly Ala Asp Trp Met Thr Cys Ile Cys Cys Ala Thr Ile Pro Thr
      85      90      95
Met Glu Ala Ala Leu Lys Ala Ile Lys Thr Glu Arg Gly Glu Arg Gly
45      100      105      110
Glu Ile Gln Ile Glu Leu Tyr Gly Asp Trp Thr Phe Glu Gln Ala Gln
      115      120      125
Leu Trp Leu Asp Ala Gly Ile Ser Gln Ala Ile Tyr His Gln Ser Arg
50      130      135      140
Asp Ala Leu Leu Ala Gly Glu Thr Trp Gly Glu Lys Asp Leu Asn Lys
55      145      150      155      160
Val Lys Lys Leu Ile Asp Met Gly Phe Arg Val Ser Val Thr Gly Gly
      165      170      175
Leu Asp Val Asp Thr Leu Lys Leu Phe Glu Gly Val Asp Val Phe Thr
60      180      185      190

```

-233-

Phe Ile Ala Gly Arg Gly Ile Thr Glu Ala Ala Asp Pro Ala Gly Ala  
 195 200 205

5 Ala Arg Ala Phe Lys Asp Glu Ile Lys Arg Ile Trp Gly  
 210 215 220

## (2) INFORMATION FOR SEQ ID NO:146:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 161 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

15 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

20 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

25 Met Asn Leu Lys Gln Ala Leu Ile Asp Asn Asp Ser Ile Arg Leu Gly  
 1 5 10 15  
 Leu Glu Ala Asn Asn Trp Lys Glu Ala Val Lys Val Ala Val Asp Pro  
 20 25 30  
 30 Leu Ile Glu Ser Gly Ala Ile Leu Pro Glu Tyr Tyr Asp Ala Ile Ile  
 35 40 45  
 Glu Ser Thr Glu Glu Tyr Gly Pro Tyr Tyr Ile Leu Met Pro Gly Met  
 50 55 60  
 35 Ala Met Pro His Ala Arg Pro Glu Ala Gly Val Gln Ser Asp Ala Phe  
 65 70 75 80  
 40 Ser Leu Ile Thr Leu Gln Asn Pro Val Val Phe Ser Asp Gly Lys Glu  
 85 90 95  
 Val Ser Val Leu Leu Ala Leu Ala Ala Thr Ser Ser Lys Ile His Thr  
 100 105 110  
 45 Ser Val Ala Ile Pro Gln Ile Ile Ala Leu Phe Glu Leu Glu Asp Ser  
 115 120 125  
 Ile Ala Arg Leu Gln Ala Cys Gln Thr Lys Glu Asp Val Leu Ala Met  
 130 135 140  
 50 Ile Glu Glu Ser Lys Asp Ser Pro Tyr Leu Glu Gly Leu Asp Leu Glu  
 145 150 155 160  
 55 Ser

## (2) INFORMATION FOR SEQ ID NO:147:

60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 293 amino acids  
 (B) TYPE: amino acid

-234-

(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

5

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

15	Met	Ser	Arg	Asp	Ile	Ile	Lys	Leu	Asp	Gln	Ile	Asp	Val	Thr	Phe	His
	1				5					10					15	
	Gln	Lys	Lys	Arg	Thr	Ile	Thr	Ala	Val	Lys	Asp	Val	Thr	Ile	His	Ile
				20					25					30		
20	Gln	Glu	Gly	Asp	Ile	Tyr	Gly	Ile	Val	Gly	Tyr	Ser	Gly	Ala	Gly	Lys
			35					40					45			
	Ser	Thr	Leu	Val	Arg	Val	Ile	Asn	Leu	Leu	Gln	Lys	Pro	Ser	Ala	Gly
		50					55					60				
25	Lys	Ile	Thr	Ile	Asp	Asp	Asp	Val	Ile	Phe	Asp	Gly	Lys	Val	Thr	Leu
	65				70						75					80
	Thr	Ala	Glu	Gln	Leu	Arg	Arg	Lys	Arg	Gln	Asp	Ile	Gly	Met	Ile	Phe
					85					90					95	
30	Gln	His	Phe	Asn	Leu	Met	Ser	Gln	Lys	Thr	Ala	Glu	Glu	Asn	Val	Ala
				100					105					110		
	Phe	Ala	Leu	Lys	His	Ser	Gly	Leu	Ser	Lys	Glu	Glu	Lys	Lys	Ala	Lys
35			115					120					125			
	Val	Ala	Lys	Leu	Leu	Asp	Leu	Val	Gly	Leu	Ala	Asp	Arg	Ala	Glu	Asn
		130					135					140				
40	Tyr	Pro	Ser	Gln	Leu	Ser	Gly	Gly	Gln	Lys	Gln	Arg	Val	Ala	Ile	Ala
	145					150					155					160
	Arg	Ala	Leu	Ala	Asn	Asp	Pro	Lys	Ile	Leu	Ile	Ser	Asp	Glu	Ser	Thr
					165					170					175	
45	Ser	Ala	Leu	Asp	Pro	Lys	Thr	Thr	Lys	Gln	Ile	Leu	Ala	Leu	Leu	Gln
				180					185					190		
	Asp	Leu	Asn	Gln	Lys	Leu	Gly	Leu	Thr	Val	Val	Leu	Ile	Thr	His	Glu
50			195					200					205			
	Met	Gln	Ile	Val	Lys	Asp	Ile	Ala	Asn	Arg	Val	Ala	Val	Met	Gln	Asp
		210					215					220				
55	Gly	His	Leu	Ile	Glu	Glu	Gly	Ser	Val	Leu	Glu	Ile	Phe	Ser	Asn	Pro
	225					230					235					240
	Lys	Gln	Pro	Leu	Thr	Gln	Asp	Phe	Ile	Ser	Thr	Ala	Thr	Gly	Ile	Asp
					245					250					255	
60	Glu	Ala	Met	Val	Lys	Ile	Glu	Lys	Gln	Glu	Ile	Val	Glu	His	Leu	Ser

-235-

260 265 270

Glu Asn Ser Leu Leu Val Gln Leu Gln Val Arg Trp Ser Phe Asn Arg  
 275 280 285

5 Arg Ala Thr Phe Glu  
 290

(2) INFORMATION FOR SEQ ID NO:148:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 209 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 15 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

20 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

25 Arg Asp Val Asn Phe Glu Ile Glu Lys Gly Glu Leu Val Ile Ile Leu  
 1 5 10 15

Gly Ala Ser Gly Ala Gly Lys Ser Thr Val Leu Asn Leu Leu Gly Gly  
 20 25 30

Met Asp Thr Asn Asp Glu Gly Glu Ile Trp Ile Asp Gly Val Asn Ile  
 35 40 45

Ala Asp Tyr Ser Ser His Gln Arg Thr Asn Tyr Arg Arg Asn Asp Val  
 50 55 60

Gly Phe Val Phe Gln Phe Tyr Asn Leu Val Ser Asn Leu Thr Ala Lys  
 65 70 75 80

40 Glu Asn Val Glu Leu Ser Glu Ile Val Thr Asp Ala Leu Asn Ser Asp  
 85 90 95

Gln Val Leu Thr Asp Val Gly Leu Ala His Arg Leu Asn Asn Phe Pro  
 100 105 110

Ala Gln Leu Ser Gly Gly Glu Gln Gln Arg Val Ser Ile Ala Arg Ala  
 115 120 125

50 Val Ala Lys Asn Pro Lys Ile Leu Leu Cys Asp Glu Pro Thr Gly Ala  
 130 135 140

Leu Asp Tyr Gln Thr Gly Lys Gln Val Leu Lys Ile Leu Gln Asp Met  
 145 150 155 160

55 Ser Arg Gln Lys Gly Ala Thr Val Ile Ile Val Thr His Asn Gly Ala  
 165 170 175

Leu Ala Pro Ile Ala Asp Arg Val Ile Gln Met His Asp Ala Ser Val  
 180 185 190

60



-236-

Lys Asp Val Val Leu Asn Gln His Pro Gln Asp Ile Asp Ser Leu Glu  
 195 200 205

Tyr

5

## (2) INFORMATION FOR SEQ ID NO:149:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 213 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

10

15 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

20 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

25

Met Ile Glu Leu Lys Asn Ile Thr Lys Thr Ile Gly Gly Lys Val Ile  
 1 5 10 15

30

Leu Asp Asn Leu Ser Leu Arg Ile Asp Gln Gly Asp Leu Val Ala Ile  
 20 25 30

Val Gly Lys Ser Gly Ser Gly Lys Ser Thr Leu Leu Asn Leu Leu Gly  
 35 40 45

35

Leu Ile Asp Gly Asp Tyr Ser Gly Arg Tyr Glu Ile Phe Gly Gln Thr  
 50 55 60

40

Asn Leu Ala Val Asn Ser Ala Lys Ser Gln Thr Ile Ile Arg Glu His  
 65 70 75 80

Ile Ser Tyr Leu Phe Gln Asn Phe Ala Leu Ile Asp Asp Glu Thr Val  
 85 90 95

45

Glu Tyr Asn Leu Met Leu Ala Leu Lys Tyr Val Lys Leu Pro Lys Lys  
 100 105 110

Asp Lys Leu Lys Lys Val Glu Glu Ile Leu Glu Arg Val Gly Leu Ser  
 115 120 125

50

Ala Thr Leu His Gln Arg Val Ser Glu Leu Ser Gly Gly Glu Gln Gln  
 130 135 140

55

Arg Ile Ala Val Ala Arg Ala Ile Leu Lys Pro Ser Gln Leu Ile Leu  
 145 150 155 160

Ala Asp Glu Pro Thr Gly Ser Leu Asp Pro Glu Asn Arg Asp Leu Val  
 165 170 175

60

Leu Lys Phe Leu Leu Glu Met Asn Arg Glu Gly Lys Thr Val Ile Ile  
 180 185 190

-237-

Val Thr His Asp Ala Tyr Val Ala Gln Gln Cys His Arg Val Ile Glu  
 195 200 205

Leu Gly Glu Gly Lys  
 210

5

## (2) INFORMATION FOR SEQ ID NO:150:

## (i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 84 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

15

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

Ala Lys Pro Lys Gln Leu Ser Gly Gly Gln Lys Gln Arg Val Ala Ile  
 1 5 10 15

25

Ala Arg Ala Leu Ser Met Asn Pro Asp Ala Ile Leu Phe Asp Glu Pro  
 20 25 30

Thr Ser Ala Leu Asp Pro Glu Met Val Gly Glu Val Leu Lys Ile Met  
 35 40 45

30

Gln Asp Leu Ala Gln Glu Gly Leu Thr Met Ile Val Val Thr His Glu  
 50 55 60

35

Met Glu Phe Ala Arg Asp Val Ser His Arg Val Ile Phe Met Asp Lys  
 65 70 75 80

Gly Val Ile Pro

40

## (2) INFORMATION FOR SEQ ID NO:151:

## (i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 118 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

50

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

Tyr Tyr Gly Asp Tyr His Ala Leu Arg Asn Ile Asn Leu Arg Phe Glu  
 1 5 10 15

60

-238-

Lys Gly Gln Val Val Val Leu Leu Gly Pro Ser Gly Ser Gly Lys Ser  
 20 25 30  
 5 Thr Leu Ile Arg Thr Ile Asn Gly Leu Glu Ala Val Asp Lys Gly Ser  
 35 40 45  
 Leu Leu Val Asn Gly His Gln Val Ala Gly Ala Ser Gln Lys Asp Leu  
 50 55 60  
 10 Val Pro Leu Arg Lys Glu Val Gly Met Val Phe Gln His Phe Asn Leu  
 65 70 75 80  
 Tyr Pro His Lys Thr Val Leu Glu Asn Val Thr Leu Ala Pro Ile Lys  
 85 90 95  
 15 Val Leu Gly Ile Asp Lys Lys Glu Ala Glu Lys Thr Ala Gln Lys Tyr  
 100 105 110  
 Leu Glu Phe Val Asn Met  
 115  
 20

## (2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:  
 25 (A) LENGTH: 237 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 30 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO  
 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:  
 40 Met Thr Lys Lys Gln Leu His Leu Val Ile Val Thr Gly Met Gly Gly  
 1 5 10 15  
 Ala Gly Lys Thr Val Ala Ile Gln Ser Phe Glu Asp Leu Gly Tyr Phe  
 20 25 30  
 45 Thr Ile Asp Asn Met Pro Pro Ala Leu Leu Pro Lys Phe Leu Gln Leu  
 35 40 45  
 Val Glu Ile Lys Glu Asp Asn Pro Lys Leu Ala Leu Val Val Asp Met  
 50 55 60  
 Arg Ser Arg Ser Phe Phe Ser Glu Ile Gln Ala Val Leu Asp Glu Leu  
 65 70 75 80  
 55 Glu Asn Gln Asp Gly Leu Asp Phe Lys Ile Leu Phe Leu Asp Ala Ala  
 85 90 95  
 Asp Lys Glu Leu Val Ala Arg Tyr Lys Glu Thr Arg Arg Ser His Pro  
 100 105 110  
 60 Leu Ala Ala Asp Gly Arg Ile Leu Asp Gly Ile Lys Leu Glu Arg Glu  
 115 120 125

-239-

Leu Leu Ala Pro Leu Lys Asn Met Ser Gln Asn Val Val Asp Thr Thr  
 130 135 140  
 5 Glu Leu Thr Pro Arg Glu Leu Arg Lys Thr Leu Ala Glu Gln Phe Ser  
 145 150 155 160  
 Asp Gln Glu Gln Ala Gln Ser Phe Arg Ile Glu Val Met Ser Phe Gly  
 165 170 175  
 10 Phe Lys Tyr Gly Ile Pro Ile Asp Ala Asp Leu Val Phe Asp Val Arg  
 180 185 190  
 Phe Leu Pro Asn Pro Tyr Tyr Leu Pro Glu Leu Arg Asn Gln Thr Gly  
 195 200 205  
 Val Asp Glu Pro Val Tyr Asp Tyr Val Met Asn His Pro Glu Ser Glu  
 210 215 220  
 20 Asp Phe Tyr Gln His Leu Leu Ala Leu Ile Glu Pro Ile  
 225 230 235

## (2) INFORMATION FOR SEQ ID NO:153:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 180 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

30 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

35 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

40 Met Leu Glu Asn Asp Ile Lys Lys Val Leu Val Ser His Asp Glu Ile  
 1 5 10 15  
 Thr Glu Ala Ala Lys Lys Leu Gly Ala Gln Leu Thr Lys Asp Tyr Ala  
 20 25 30  
 45 Gly Lys Asn Pro Ile Leu Val Gly Ile Leu Lys Gly Ser Ile Pro Phe  
 35 40 45  
 Met Ala Glu Leu Val Lys His Ile Asp Thr His Ile Glu Met Asp Phe  
 50 50 55 60  
 Met Met Val Ser Ser Tyr His Gly Gly Thr Ala Ser Ser Gly Val Ile  
 65 70 75 80  
 55 Asn Ile Lys Gln Asp Val Thr Gln Asp Ile Lys Gly Arg His Val Leu  
 85 90 95  
 Phe Val Glu Asp Ile Ile Asp Thr Gly Gln Thr Leu Lys Asn Leu Arg  
 100 105 110  
 60 Asp Met Phe Lys Glu Arg Glu Ala Ala Ser Val Lys Ile Ala Thr Leu

-240-

115                      120                      125  
 Leu Asp Lys Pro Glu Gly Arg Val Val Glu Ile Glu Ala Asp Tyr Thr  
       130                      135                      140  
 5  
 Cys Phe Thr Ile Pro Asn Glu Phe Val Val Gly Tyr Gly Leu Asp Tyr  
       145                      150                      155                      160  
 10  
 Lys Glu Asn Tyr Arg Asn Leu Pro Tyr Ile Gly Val Leu Lys Glu Glu  
                                  165                      170                      175  
 Val Tyr Ser Asn  
                                  180  
 15    (2) INFORMATION FOR SEQ ID NO:154:  
       (i) SEQUENCE CHARACTERISTICS:  
           (A) LENGTH: 193 amino acids  
           (B) TYPE: amino acid  
 20        (C) STRANDEDNESS: not relevant  
           (D) TOPOLOGY: not relevant  
       (ii) MOLECULE TYPE: peptide  
 25    (iii) HYPOTHETICAL: NO  
       (iv) ANTI-SENSE: NO  
 30    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:  
 Met Lys Ile Gly Ile Leu Ala Leu Gln Gly Ala Phe Ala Glu His Ala  
   1                      5                      10                      15  
 35    Lys Val Leu Asp Gln Leu Gly Val Glu Ser Val Glu Leu Arg Asn Leu  
                                  20                      25                      30  
 Asp Asp Phe Gln Gln Asp Gln Ser Asp Leu Ser Gly Leu Ile Leu Pro  
                                  35                      40                      45  
 40    Gly Gly Glu Ser Thr Thr Met Gly Lys Leu Leu Arg Asp Gln Asn Met  
                                  50                      55                      60  
 45    Leu Leu Pro Ile Arg Glu Ala Ile Leu Ser Gly Leu Pro Val Phe Gly  
                                  65                      70                      75                      80  
 Thr Cys Ala Gly Leu Ile Leu Leu Ala Lys Glu Ile Thr Ser Gln Lys  
                                  85                      90                      95  
 50    Glu Ser His Leu Gly Thr Met Asp Met Val Val Glu Arg Asn Ala Tyr  
                                  100                      105                      110  
 Gly Arg Gln Leu Gly Ser Phe Tyr Thr Glu Ala Glu Cys Lys Gly Val  
                                  115                      120                      125  
 55    Gly Lys Ile Pro Met Thr Phe Ile Arg Gly Pro Ile Ile Ser Ser Val  
                                  130                      135                      140  
 60    Gly Glu Gly Val Glu Ile Leu Ala Ile Val Asn Asn Gln Ile Val Ala  
                                  145                      150                      155                      160

Ala Gln Glu Lys Asn Met Leu Val Ser Ser Phe His Pro Glu Leu Thr  
165 170 175

Ser

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 189 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: not relevant

20 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

Glu Ser Glu Val Leu Ser Pro Ala Asp Asp Arg Phe His Val Asp Lys  
1 5 10 15

Lys Glu Phe Gln Val Pro Phe Val Cys Gly Ala Lys Asp Leu Gly Glu  
20 25 30

Ala Leu Arg Arg Ile Ala Glu Gly Ala Ser Met Ile Arg Thr Lys Gly  
35 40 45

Glu Pro Gly Thr Gly Asp Ile Val Gln Ala Val Arg His Met Arg Met  
50 55 60

Met Asn Gln Glu Ile Arg Arg Ile Gln Asn Leu Arg Glu Asp Glu Leu  
65 70 75 80

Tyr Val Ala Ala Lys Asp Leu Gln Val Pro Val Glu Leu Val Gln Tyr  
85 90 95

Val	His	Glu	His	Gly	Lys	Leu	Pro	Val	Val	Asn	Phe	Ala	Ala	Gly	Gly
			100				105						110		

Val Ala Thr Pro Ala Asp Ala Ala Leu Met Met Gln Leu Gly Ala Glu  
115 120 125

Gly Val Phe Val Gly Ser Gly Ile Phe Lys Ser Gly Asp Pro Val Lys  
130 135 140

Arg Ala Ser Ala Ile Val Lys Ala Val Thr Asn Phe Arg Asn Pro Gln  
145 150 155 160

Ile Leu Ala Gln Ile Ser Glu Asp Leu Gly Glu Ala Met Val Gly Ile  
165 170 175

Asn Glu Asn Ile Gln Ile Leu Met Ala Glu Arg Gly Lys

-242-

180

185

## (2) INFORMATION FOR SEQ ID NO:156:

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 162 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

10 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

Asp	Lys	Gly	Trp	Phe	Val	Leu	Gln	Thr	Tyr	Ser	Gly	Tyr	Glu	Asn	Lys
1				5					10					15	
Val	Lys	Glu	Asn	Leu	Leu	Gln	Arg	Ala	Gln	Thr	Tyr	Asn	Met	Leu	Asp
			20					25					30		
Asn	Ile	Leu	Arg	Val	Glu	Ile	Pro	Thr	Gln	Thr	Val	Gln	Val	Glu	Lys
		35					40					45			
Asn	Gly	Lys	Arg	Lys	Glu	Val	Glu	Glu	Asn	Arg	Phe	Pro	Gly	Tyr	Val
	50					55					60				
Leu	Val	Glu	Met	Val	Met	Thr	Asp	Glu	Ala	Trp	Phe	Val	Val	Arg	Asn
65					70					75					80
Ala	Gln	Ser	Pro	Thr	Lys	Phe	Ile	Ser	Glu	Gln	Thr	Ala	Tyr	Glu	Ile
				85					90					95	
Asp	Glu	Glu	Val	Arg	Ser	Leu	Leu	Asn	Glu	Ala	Arg	Asn	Lys	Ala	Ala
			100					105					110		
Glu	Ile	Ile	Gln	Ser	Asn	Arg	Glu	Thr	His	Lys	Leu	Ile	Ala	Glu	Ala
		115					120					125			
Leu	Leu	Lys	Tyr	Glu	Thr	Leu	Asp	Ser	Thr	Gln	Ile	Lys	Ala	Leu	Tyr
	130					135					140				
Glu	Thr	Gly	Lys	Met	Pro	Glu	Ser	Ser	Arg	Arg	Gly	Ile	Ser	Cys	Thr
145					150					155					160
Ile	Leu														

## 55 (2) INFORMATION FOR SEQ ID NO:157:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 208 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

60

-243-

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

Val Asn Ser Ser Ser Val Pro Gly Asp Arg Phe Ser Val Leu Leu Glu  
 1 5 10 15  
 His Lys Gly Ile His Pro Ile Val Tyr Ile Ser Lys Met Asp Leu Leu  
 20 25 30  
 Glu Asp Arg Gly Glu Leu Asp Phe Tyr Gln Gln Thr Tyr Gly Asp Ile  
 35 40 45  
 Gly Tyr Asp Phe Val Thr Ser Lys Glu Glu Leu Leu Ser Leu Leu Thr  
 50 55 60  
 Gly Lys Val Thr Val Phe Met Gly Gln Thr Gly Val Gly Lys Ser Thr  
 65 70 75 80  
 Leu Leu Asn Lys Ile Ala Pro Asp Leu Asn Leu Glu Thr Gly Glu Ile  
 85 90 95  
 Ser Asp Ser Leu Gly Arg Gly Arg His Thr Thr Arg Ala Val Ser Phe  
 100 105 110  
 Tyr Asn Leu Asn Gly Gly Lys Ile Ala Asp Thr Pro Gly Phe Ser Ser  
 115 120 125  
 Leu Asp Tyr Glu Val Ser Arg Ala Glu Asp Leu Asn Gln Ala Phe Pro  
 130 135 140  
 Glu Ile Ala Thr Val Ser Arg Asp Cys Lys Phe Arg Thr Cys Thr His  
 145 150 155 160  
 Thr His Glu Pro Ser Cys Ala Val Lys Pro Ala Val Glu Glu Gly Val  
 165 170 175  
 Ile Ala Thr Phe Arg Phe Asp Asn Tyr Leu Gln Phe Leu Ser Glu Ile  
 180 185 190  
 Glu Asn Arg Arg Glu Thr Tyr Lys Lys Val Ser Lys Lys Ile Pro Lys  
 195 200 205

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 403 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide



-244-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

10	Gln	Gln	Ser	Val	Lys	Lys	Lys	Val	Leu	Pro	Ala	Ile	Glu	Arg	Arg	Ile	1	5	10	15
	Arg	Thr	Glu	Leu	Thr	Glu	Lys	Ala	Glu	Gly	Ala	Ile	Gln	Leu	Phe		20	25	30	
15	Ser	Asp	Asn	Leu	Arg	Asn	Leu	Leu	Val	Ala	Pro	Leu	Lys	Gly	Arg		35	40	45	
	Val	Val	Leu	Gly	Phe	Asp	Pro	Ala	Phe	Arg	Thr	Gly	Ala	Lys	Leu	Ala	50	55	60	
20	Val	Val	Asp	Ala	Thr	Gly	Lys	Met	Leu	Thr	Thr	Gln	Val	Ile	Tyr	Pro	65	70	75	80
	Val	Lys	Pro	Ala	Ser	Ala	Arg	Gln	Ile	Glu	Glu	Ala	Lys	Lys	Asp	Leu	85	90	95	
25	Ala	Asp	Leu	Ile	Gly	Gln	Tyr	Gly	Val	Glu	Ile	Ile	Ala	Ile	Gly	Asn	100	105	110	
30	Gly	Thr	Ala	Ser	Arg	Glu	Ser	Glu	Ala	Phe	Val	Ala	Glu	Val	Leu	Lys	115	120	125	
	Asp	Phe	Pro	Glu	Val	Ser	Tyr	Val	Ile	Val	Asn	Glu	Ser	Gly	Ala	Ser	130	135	140	
35	Val	Tyr	Ser	Ala	Ser	Glu	Leu	Ala	Arg	Gln	Glu	Phe	Pro	Asp	Leu	Thr	145	150	155	160
	Val	Glu	Lys	Arg	Ser	Ala	Ile	Ser	Ile	Ala	Arg	Arg	Leu	Gln	Asp	Pro	165	170	175	
40	Leu	Ala	Glu	Leu	Val	Lys	Ile	Asp	Pro	Lys	Ser	Ile	Gly	Val	Gly	Gln	180	185	190	
45	Tyr	Gln	His	Asp	Val	Ser	Gln	Lys	Lys	Leu	Ser	Glu	Ser	Leu	Asp	Phe	195	200	205	
	Val	Val	Asp	Thr	Val	Val	Asn	Gln	Val	Gly	Val	Asn	Val	Asn	Thr	Ala	210	215	220	
50	Ser	Pro	Ala	Leu	Leu	Ser	His	Val	Ala	Gly	Leu	Asn	Lys	Thr	Ile	Ser	225	230	235	240
	Glu	Asn	Ile	Val	Lys	Tyr	Arg	Glu	Glu	Glu	Gly	Lys	Ile	Thr	Ser	Arg	245	250	255	
55	Ala	Gln	Ile	Lys	Lys	Val	Pro	Arg	Leu	Gly	Ala	Lys	Ala	Phe	Glu	Gln	260	265	270	
60	Ala	Ala	Gly	Phe	Leu	Arg	Ile	Pro	Glu	Ser	Ser	Asn	Ile	Leu	Asp	Asn	275	280	285	

-245-

Thr Gly Val His Pro Glu Asn Tyr Thr Ala Val Lys Leu Phe Lys Arg  
 290 295 300  
 5 Leu Asp Ile Lys Asp Leu Asn Glu Glu Ala Ser Lys Leu Lys Ser Leu  
 305 310 315 320  
 Ser Val Lys Glu Met Ala Gln Glu Leu Asp Leu Gly Pro Glu Thr Leu  
 325 330 335  
 10 Lys Asp Ile Ile Ala Asp Leu Leu Lys Pro Gly Arg Asp Phe Arg Asp  
 340 345 350  
 Ser Phe Asp Ala Pro Val Leu Arg Gln Asp Val Leu Asp Ile Lys Asp  
 355 360 365  
 15 Leu Val Val Gly Gln Lys Leu Glu Gly Val Val Arg Asn Val Val Asp  
 370 375 380  
 20 Phe Gly Ala Phe Val Asp Ile Gly Val His Glu Asp Gly Leu Ile His  
 385 390 395 400  
 Ile Leu Ile

25

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 179 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

30

(ii) MOLECULE TYPE: peptide

35

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

45 Met Phe Arg Ala Ala Met Ala Asn Gln Thr Glu Met Gly Val Leu Ala  
 1 5 10 15  
 Lys Ser Tyr Ile Asp Lys Gly Glu Leu Val Pro Asp Glu Val Thr Asn  
 20 25 30  
 50 Gly Ile Val Lys Glu Arg Leu Ser Gln Asp Asp Ile Lys Glu Thr Gly  
 35 40 45  
 55 Phe Leu Leu Asp Gly Tyr Pro Arg Thr Ile Glu Gln Ala His Ala Leu  
 50 55 60  
 Asp Lys Thr Leu Ala Glu Leu Gly Ile Glu Leu Glu Gly Ile Ile Asn  
 65 70 75 80  
 60 Ile Glu Val Asn Pro Asp Ser Leu Leu Glu Arg Leu Ser Gly Arg Ile  
 85 90 95

-246-

Ile His Arg Val Thr Gly Glu Thr Phe His Lys Val Phe Asn Pro Pro  
                   100                  105                  110

5 Val Asp Tyr Lys Glu Glu Asp Tyr Tyr Gln Arg Glu Asp Asp Lys Pro  
                   115                  120                  125

Glu Thr Val Lys Arg Arg Leu Asp Val Asn Ile Ala Gln Gly Glu Pro  
           130                  135                  140

10 Ile Ile Ala His Tyr Arg Ala Lys Gly Leu Val His Asp Ile Glu Gly  
      145                  150                  155                  160

15 Asn Gln Asp Ile Asn Asp Val Phe Ser Asp Ile Glu Lys Val Leu Thr  
                   165                  170                  175

Asn Leu Lys

20 (2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:  
       (A) LENGTH: 191 amino acids  
       (B) TYPE: amino acid  
       (C) STRANDEDNESS: not relevant  
       (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

30 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

40 Met Ile Glu Phe Glu Lys Pro Asn Ile Thr Lys Ile Asp Glu Asn Lys  
      1                  5                  10                  15

Asp Tyr Gly Lys Phe Val Ile Glu Pro Leu Glu Arg Gly Tyr Gly Thr  
           20                  25                  30

45 Thr Leu Gly Asn Ser Leu Arg Arg Val Leu Leu Ala Ser Leu Pro Gly  
           35                  40                  45

Ala Ala Val Thr Ser Ile Asn Ile Asp Gly Val Leu His Glu Phe Asp  
      50                  55                  60

50 Thr Val Pro Gly Val Arg Glu Asp Val Met Gln Ile Ile Leu Asn Ile  
      65                  70                  75                  80

55 Lys Gly Ile Ala Val Lys Ser Tyr Val Glu Asp Glu Lys Ile Ile Glu  
           85                  90                  95

Leu Asp Val Glu Gly Pro Ala Glu Val Thr Ala Gly Asp Ile Leu Thr  
          100                  105                  110

60 Asp Ser Asp Ile Glu Ile Val Asn Pro Asp His Tyr Leu Phe Thr Ile  
          115                  120                  125

.-247-

Gly Glu Gly Ser Ser Leu Lys Ala Thr Met Thr Val Asn Ser Gly Arg  
 130 135 140

5 Gly Tyr Val Pro Ala Asp Glu Asn Lys Lys Asp Asn Ala Pro Val Gly  
 145 150 155 160

Thr Leu Ala Val Asp Ser Ile Tyr Thr Pro Val Thr Lys Val Asn Tyr  
 165 170 175

10 Gln Val Glu Pro Ala Arg Val Gly Ser Asn Asp Gly Phe Asp Ser  
 180 185 190

(2) INFORMATION FOR SEQ ID NO:161:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 335 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 20 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

30 Glu Tyr Leu Gly Ala Thr Val Gln Val Ile Pro His Ile Thr Asp Ala  
 1 5 10 15

35 Leu Lys Glu Lys Ile Lys Ser Ala Ala Leu Thr Thr Asp Ser Asp Val  
 20 25 30

Ile Ile Thr Glu Val Gly Gly Thr Val Gly Asp Ile Glu Ser Leu Pro  
 35 40 45

40 Phe Leu Glu Ala Leu Arg Gln Met Lys Ala Asp Val Gly Ala Asp Asn  
 50 55 60

Val Met Tyr Ile His Thr Thr Leu Pro Tyr Leu Lys Ala Ala Gly Glu  
 65 70 75 80

45 Met Lys Lys Pro Thr Gln His Ser Val Lys Leu Arg Gly Leu Gly Ile  
 85 90 95

50 Gln Pro Asn Met Leu Val Ile Arg Thr Glu Glu Pro Ala Gly Gln Gly  
 100 105 110

Ile Lys Asn Lys Leu Ala Gln Phe Cys Asp Val Ala Pro Glu Ser Leu  
 115 120 125

55 Ile Glu Ser Leu Asp Val Glu His Leu Tyr Gln Ile Pro Leu Asn Leu  
 130 135 140

Gln Ala Gln Gly Met Asp Gln Ile Val Cys Asp His Leu Lys Leu Asp  
 145 150 155 160

60 Ala Pro Ala Ala Asp Met Thr Glu Trp Ser Ala Met Val Asp Lys Val

-248-

	165	170	175
5	Met Asn Leu Lys Lys Gln Val Lys Ile Ser Leu Val Gly Lys Tyr Val 180 185 190		
	Glu Leu Gln Asp Ala Tyr Ile Ser Val Val Glu Ala Leu Lys His Ser 195 200 205		
10	Gly Tyr Val Asn Asp Val Glu Val Lys Ile Asn Trp Val Asn Ala Asn 210 215 220		
	Asp Val Thr Ala Glu Asn Val Ala Glu Leu Leu Ser Asp Ala Asp Gly 225 230 235 240		
15	Ile Ile Val Pro Gly Gly Phe Gly Gln Arg Gly Thr Glu Gly Lys Ile 245 250 255		
	Gln Ala Ile Arg Tyr Ala Arg Glu Asn Asp Val Pro Met Leu Gly Val 260 265 270		
20	Cys Leu Gly Met Gln Leu Thr Cys Ile Glu Phe Ala Arg His Val Leu 275 280 285		
	Gly Leu Glu Gly Ala Asn Ser Ala Glu Leu Ala Pro Glu Thr Lys Tyr 290 295 300		
25	Pro Ile Ile Asp Ile Met Arg Asp Gln Ile Asp Ile Glu Asp Met Gly 305 310 315 320		
	Gly Thr Leu Arg Leu Gly Leu Tyr Pro Ser Lys Leu Lys Arg Leu 325 330 335		

## (2) INFORMATION FOR SEQ ID NO:162:

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 301 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

40 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

45 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

50	Met Ser Glu Lys Leu Val Glu Ile Lys Asp Leu Glu Ile Ser Phe Gly 1 5 10 15
	Glu Gly Ser Lys Lys Phe Val Ala Val Lys Asn Ala Asn Phe Phe Ile 20 25 30
55	Asn Lys Gly Glu Thr Phe Ser Leu Val Gly Glu Ser Gly Ser Gly Lys 35 40 45
60	Thr Thr Ile Gly Arg Ala Ile Ile Gly Leu Asn Asp Thr Ser Asn Gly 50 55 60

-249-

Asp Ile Ile Phe Asp Gly Gln Lys Ile Asn Gly Lys Lys Ser Arg Glu  
 65 70 75 80  
 5 Gln Ala Ala Glu Leu Ile Arg Arg Ile Gln Met Ile Phe Gln Asp Pro  
 85 90 95  
 Ala Ala Ser Leu Asn Glu Arg Ala Thr Val Asp Tyr Ile Ile Ser Glu  
 100 105 110  
 10 Gly Leu Tyr Asn His Arg Leu Phe Lys Asp Glu Glu Glu Arg Lys Glu  
 115 120 125  
 Lys Val Gln Ser Ile Ile Arg Glu Val Gly Leu Leu Ala Glu His Leu  
 130 135 140  
 15 Thr Arg Tyr Pro His Glu Phe Ser Gly Gly Gln Arg Gln Arg Ile Gly  
 145 150 155 160  
 Ile Ala Arg Ala Leu Val Met Gln Pro Asp Phe Val Ile Ala Asp Glu  
 165 170 175  
 20 Pro Ile Ser Ala Leu Asp Val Ser Val Arg Ala Gln Val Leu Asn Leu  
 180 185 190  
 25 Leu Lys Lys Phe Gln Lys Glu Leu Gly Leu Thr Tyr Leu Phe Ile Ala  
 195 200 205  
 His Asp Leu Ser Val Val Arg Phe Ile Ser Asp Arg Ile Ala Val Ile  
 210 215 220  
 30 Tyr Lys Gly Val Ile Val Glu Val Ala Glu Thr Glu Glu Leu Phe Asn  
 225 230 235 240  
 Asn Pro Ile His Pro Tyr Thr Gln Ala Leu Leu Ser Ala Val Pro Ile  
 245 250 255  
 35 Pro Asp Pro Ile Leu Glu Arg Lys Lys Val Leu Lys Val Tyr Asp Pro  
 260 265 270  
 40 Ser Gln His Asp Tyr Glu Thr Asp Lys Pro Ser Met Val Glu Ile Arg  
 275 280 285  
 Pro Gly His Tyr Val Trp Ala Asn Gln Ala Glu Leu Ala  
 290 295 300  
 45

## (2) INFORMATION FOR SEQ ID NO:163:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 151 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO

60

-250-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

5      Gln Ile Gln Lys Ser Phe Lys Gly Gln Ser Pro Tyr Gly Lys Leu Tyr  
       1                    5                    10                    15  
       Leu Val Ala Thr Pro Ile Gly Asn Leu Asp Asp Met Thr Phe Arg Ala  
                   20                    25                    30  
 10      Ile Gln Thr Leu Lys Glu Val Asp Trp Ile Ala Ala Glu Asp Thr Arg  
                   35                    40                    45  
       Asn Thr Gly Leu Leu Leu Lys His Phe Asp Ile Ser Thr Lys Gln Ile  
                   50                    55                    60  
 15      Ser Phe His Glu His Asn Ala Lys Glu Lys Ile Pro Asp Leu Ile Gly  
       65                    70                    75                    80  
       Phe Leu Lys Ala Gly Gln Ser Ile Ala Gln Val Ser Asp Ala Gly Leu  
                   85                    90                    95  
 20      Pro Ser Ile Ser Asp Pro Gly His Asp Leu Val Lys Ala Ala Ile Glu  
                   100                    105                    110  
 25      Glu Glu Ile Ala Val Val Thr Val Pro Gly Ala Ser Ala Gly Ile Ser  
                   115                    120                    125  
       Ala Leu Ile Ala Ser Gly Leu Ala Pro Gln Pro His Ile Phe Tyr Gly  
                   130                    135                    140  
 30      Phe Leu Pro Arg Lys Ser Gly  
       145                    150

## (2) INFORMATION FOR SEQ ID NO:164:

35      (i) SEQUENCE CHARACTERISTICS:  
           (A) LENGTH: 258 amino acids  
           (B) TYPE: amino acid  
 40      (C) STRANDEDNESS: not relevant  
           (D) TOPOLOGY: not relevant  
       (ii) MOLECULE TYPE: peptide  
       (iii) HYPOTHETICAL: NO  
 45      (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

55      Ser Arg Lys Asp Lys Gln Glu Arg Ile Ser Lys Glu Thr Met Glu Ile  
       1                    5                    10                    15  
       Tyr Ala Pro Leu Ala His Arg Leu Gly Ile Ser Ser Val Lys Trp Glu  
                   20                    25                    30  
 60      Leu Glu Asp Leu Ser Phe Arg Tyr Leu Asn Pro Thr Glu Phe Tyr Lys  
                   35                    40                    45

-251-

Ile Thr His Met Met Lys Glu Lys Arg Arg Glu Arg Glu Ala Leu Val  
 50 55 60  
 5 Asp Glu Val Val Thr Lys Leu Glu Glu Tyr Thr Thr Glu Arg His Leu  
 65 70 75 80  
 Lys Gly Lys Ile Tyr Gly Arg Pro Lys His Ile Tyr Ser Ile Phe Arg  
 85 90 95  
 10 Lys Met Gln Asp Lys Arg Lys Arg Phe Glu Glu Ile Tyr Asp Leu Ile  
 100 105 110  
 Ala Ile Arg Cys Ile Leu Asp Thr Gln Ser Asp Val Tyr Ala Met Leu  
 115 120 125  
 15 Gly Tyr Val His Glu Phe Trp Lys Pro Met Pro Gly Arg Phe Lys Asp  
 130 135 140  
 Tyr Ile Ala Asn Arg Lys Ala Asn Gly Tyr Gln Ser Ile His Thr Thr  
 145 150 155 160  
 Val Tyr Gly Pro Lys Gly Pro Ile Glu Phe Gln Ile Arg Thr Lys Glu  
 165 170 175  
 25 Met His Glu Val Ala Glu Tyr Gly Val Ala Ala His Trp Ala Tyr Lys  
 180 185 190  
 Lys Gly Ile Lys Gly Gln Val Asn Ser Lys Glu Ser Ala Ile Gly Met  
 195 200 205  
 30 Asn Trp Ile Lys Glu Met Met Glu Leu Gln Asp Gln Ala Asp Asp Ala  
 210 215 220  
 Lys Glu Phe Val Asp Ser Val Lys Glu Asn Tyr Leu Ala Glu Glu Ile  
 225 230 235 240  
 Thr Val Leu Pro Gln Met Glu Leu Ser Val Pro Ser Gln Arg Phe Arg  
 245 250 255  
 40 Thr Asp

## (2) INFORMATION FOR SEQ ID NO:165:

- 45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 289 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 50 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 55 (iv) ANTI-SENSE: NO  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:  
 60 Thr Lys Val Gly Gly Glu Ala Asp Tyr Leu Val Phe Pro Arg Asn Arg  
 1 5 10 15



-252-

Phe Glu Leu Ala Arg Val Val Lys Phe Ala Asn Gln Glu Asn Ile Pro  
                     20                    25                    30  
 5 Trp Met Val Leu Gly Asn Ala Ser Asn Ile Ile Val Arg Asp Gly Gly  
                     35                    40                    45  
 Ile Arg Gly Phe Val Ile Leu Cys Asp Lys Leu Asn Asn Val Ser Val  
                     50                    55                    60  
 10 Asp Gly Tyr Thr Ile Glu Ala Glu Ala Gly Ala Asn Leu Ile Glu Thr  
                     65                    70                    75                    80  
 Thr Arg Ile Ala Leu Arg His Ser Leu Thr Gly Phe Glu Phe Ala Cys  
                     85                    90                    95  
 Gly Ile Pro Gly Ser Val Gly Gly Ala Val Phe Met Asn Ala Gly Ala  
                     100                    105                    110  
 20 Tyr Gly Gly Glu Ile Ala His Ile Leu Gln Ser Cys Lys Val Leu Thr  
                     115                    120                    125  
 Lys Asp Gly Glu Ile Glu Thr Leu Ser Ala Lys Asp Leu Ala Phe Gly  
                     130                    135                    140  
 25 Tyr Arg His Ser Ala Ile Gln Glu Ser Gly Ala Val Val Leu Ser Val  
                     145                    150                    155                    160  
 Lys Phe Ala Leu Ala Pro Gly Thr His Gln Val Ile Lys Gln Glu Met  
                     165                    170                    175  
 Asp Arg Leu Thr His Leu Arg Glu Leu Lys Gln Pro Leu Glu Tyr Pro  
                     180                    185                    190  
 35 Ser Cys Gly Ser Val Phe Lys Arg Pro Val Gly His Phe Ala Gly Gln  
                     195                    200                    205  
 Leu Ile Ser Glu Ala Gly Leu Lys Gly Tyr Arg Ile Gly Gly Val Glu  
                     210                    215                    220  
 40 Val Ser Glu Lys His Ala Gly Phe Met Ile Asn Val Ala Asp Gly Thr  
                     225                    230                    235                    240  
 Ala Lys Asp Tyr Glu Asp Leu Ile Gln Ser Val Ile Glu Lys Val Lys  
                     245                    250                    255  
 Glu His Ser Gly Ile Thr Leu Glu Arg Glu Val Arg Ile Leu Gly Glu  
                     260                    265                    270  
 50 Ser Leu Ser Val Ala Lys Met Tyr Ala Gly Gly Phe Thr Pro Cys Lys  
                     275                    280                    285  
 Arg

55

(2) INFORMATION FOR SEQ ID NO:166:

- 60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 114 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant

-253-

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

Ala Lys Arg Arg Lys Leu Val Lys Ser Thr Thr Leu Leu Leu Ala Cys  
 1 5 10 15

15 Leu Gln Lys Pro Phe Leu Thr Thr Leu Leu Pro Thr Ile Trp Ile Cys  
 20 25 30

Val Lys Ser Ser Met Phe Thr Leu Leu Arg Leu Asn Thr Trp Ile Lys  
 35 40 45

20 Asp Phe His Ser Pro Ser Ser Cys Val Val Thr Phe Gln Lys Ala Phe  
 50 55 60

25 Thr Asn Gly Arg Gly Lys Ile Asn Lys Arg His Val Thr Cys Pro Ser  
 65 70 75 80

Phe Val Thr Met Pro Leu Thr Arg Glu Ser Ser Leu Ser Thr Thr Ser  
 85 90 95

30 Val Pro Leu Gln Met Thr Val Glu Lys Ser Ala Pro Thr Asn Val Lys  
 100 105 110

Ala Val

35

(2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 253 amino acids

40 (B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

45

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

Met Leu Lys Gln Glu Lys Leu Ala Lys Ile Leu Glu Ile Val Asn Ser  
 1 5 10 15

55

Lys Gly Thr Ile Thr Val Lys Gln Ile Met Asp Glu Ile Ala Val Ser  
 20 25 30

60

Asp Met Thr Ala Arg Arg Tyr Leu Gln Glu Leu Ala Asp Lys Asp Leu  
 35 40 45

-254-

Leu Ile Arg Val His Gly Gly Ala Glu Lys Leu Arg Thr Asn Ser Leu  
 50 55 60  
 5 Leu Thr Asn Glu Arg Ser Asn Ile Glu Lys Gln Ala Leu Gln Thr Ala  
 65 70 75 80  
 Glu Lys Gln Glu Ile Ala His Phe Ala Gly Ser Leu Val Glu Glu Arg  
 85 90 95  
 10 Glu Thr Ile Phe Ile Gly Pro Gly Thr Thr Leu Glu Phe Phe Ala Arg  
 100 105 110  
 Glu Leu Pro Ile Asp Asn Ile Arg Val Val Thr Asn Ser Leu Pro Val  
 115 120 125  
 15 Phe Leu Ile Leu Ser Glu Arg Lys Leu Thr Asp Leu Ile Leu Ile Gly  
 130 135 140  
 Gly Asn Tyr Arg Asp Ile Thr Gly Ala Phe Val Gly Thr Leu Thr Leu  
 145 150 155 160  
 Gln Asn Leu Ser Asn Leu Gln Phe Ser Lys Ala Phe Val Ser Cys Asn  
 165 170 175  
 25 Gly Ile Gln Asn Gly Ala Leu Ala Thr Phe Ser Glu Glu Glu Gly Glu  
 180 185 190  
 Ala Gln Arg Ile Ala Leu Asn Asn Ser Asn Lys Lys Tyr Leu Leu Ala  
 195 200 205  
 30 Asp His Ser Lys Phe Asn Lys Phe Asp Phe Tyr Thr Phe Tyr Asn Ile  
 210 215 220  
 Ser Asn Leu Asp Thr Ile Val Ser Asp Ser Lys Leu Ser Asp Ser Ile  
 225 230 235 240  
 Leu Phe Lys Leu Ser Lys His Ile Lys Val Ile Lys Pro  
 245 250

40 (2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 320 amino acids  
 (B) TYPE: amino acid  
 45 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

50 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

Met Glu Thr Tyr Tyr Lys Ala Ile Asn Trp Asn Ala Ile Glu Asp Val  
 1 5 10 15  
 60 Ile Asp Lys Ser Thr Trp Glu Lys Leu Thr Glu Gln Phe Trp Leu Asp  
 20 25 30

-255-

Thr Arg Ile Pro Leu Ser Asn Asp Leu Asp Asp Trp Arg Lys Leu Ser  
                   35                  40                  45  
 5 Asn Lys Glu Lys Asp Leu Val Gly Lys Val Phe Gly Gly Leu Thr Leu  
           50                  55                  60  
 Leu Asp Thr Met Gln Ser Glu Thr Gly Val Gln Ala Leu Arg Ala Asp  
 65                  70                  75                  80  
 10 Ile Arg Thr Pro His Glu Glu Ala Val Phe Asn Asn Ile Gln Phe Met  
                   85                  90                  95  
 Glu Ser Val His Ala Lys Ser Tyr Ser Ser Ile Phe Ser Thr Leu Asn  
 15                  100                  105                  110  
 Thr Lys Ala Glu Ile Glu Glu Ile Phe Glu Trp Thr Asn Thr Asn Pro  
           115                  120                  125  
 20 Tyr Leu Gln Lys Lys Ala Glu Ile Val Asn Glu Ile Tyr Leu Asn Gly  
           130                  135                  140  
 Ser Pro Leu Glu Lys Lys Val Ala Ser Val Phe Leu Glu Thr Phe Leu  
 25          145                  150                  155                  160  
 Phe Tyr Ser Gly Phe Phe Thr Pro Leu Tyr Tyr Leu Gly Asn Asn Lys  
                   165                  170                  175  
 30 Leu Ala Asn Val Ala Glu Ile Ile Lys Leu Ile Ile Arg Asp Glu Ser  
                   180                  185                  190  
 Val His Gly Thr Tyr Ile Gly Tyr Lys Phe Gln Leu Gly Phe Asn Glu  
           195                  200                  205  
 35 Leu Pro Glu Glu Glu Gln Glu Lys Leu Lys Glu Trp Met Tyr Asp Leu  
           210                  215                  220  
 Leu Tyr Thr Leu Tyr Glu Asn Glu Glu Gly Tyr Thr Glu Ser Leu Tyr  
 40          225                  230                  235                  240  
 Asp Gly Val Gly Trp Thr Glu Glu Val Lys Thr Phe Leu Arg Tyr Asn  
                   245                  250                  255  
 45 Ala Asn Lys Ala Leu Met Asn Met Gly Gln Asp Pro Leu Phe Pro Asp  
           260                  265                  270  
 Ser Ala Glu Asp Val Asn Pro Ile Val Met Asn Gly Ile Ser Thr Gly  
           275                  280                  285  
 50 Thr Ser Asn His Asp Phe Phe Ser Gln Val Gly Asn Gly Tyr Leu Leu  
           290                  295                  300  
 Gly Glu Val Glu Ala Met Gln Asp Asp Asp Tyr Asn Tyr Gly Leu Asp  
 55          305                  310                  315                  320

## (2) INFORMATION FOR SEQ ID NO:169:

## (i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 240 amino acids  
 (B) TYPE: amino acid

-256-

(C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

Ile	Glu	Glu	Gly	Val	Lys	Val	Val	Thr	Thr	Gly	Ala	Gly	Asn	Pro	Ser	1	5	10	15
Lys	Tyr	Met	Glu	Arg	Phe	His	Glu	Ala	Gly	Ile	Ile	Val	Ile	Pro	Val	20	25	30	
Val	Pro	Ser	Val	Ala	Leu	Ala	Lys	Arg	Met	Glu	Lys	Ile	Gly	Ala	Asp	35	40	45	
Ala	Val	Ile	Ala	Glu	Gly	Met	Glu	Ala	Gly	Gly	His	Ile	Gly	Lys	Leu	50	55	60	
Thr	Thr	Met	Thr	Leu	Val	Arg	Gln	Val	Ala	Thr	Ala	Val	Ser	Ile	Pro	65	70	75	80
Val	Ile	Ala	Ala	Gly	Gly	Ile	Ala	Asp	Gly	Glu	Gly	Ala	Ala	Ala	Gly	85	90	95	
Phe	Met	Leu	Gly	Ala	Glu	Ala	Val	Gln	Val	Gly	Thr	Arg	Phe	Val	Val	100	105	110	
Ala	Lys	Glu	Ser	Asn	Ala	His	Pro	Asn	Tyr	Lys	Glu	Lys	Ile	Leu	Lys	115	120	125	
Ala	Arg	Asp	Ile	Asp	Thr	Thr	Ile	Ser	Ala	Gln	His	Phe	Gly	His	Ala	130	135	140	
Val	Arg	Ala	Ile	Lys	Asn	Gln	Leu	Thr	Arg	Asp	Phe	Glu	Leu	Ala	Glu	145	150	155	160
Lys	Asp	Ala	Phe	Lys	Gln	Glu	Asp	Pro	Asp	Leu	Glu	Ile	Phe	Glu	Gln	165	170	175	
Met	Gly	Ala	Gly	Ala	Leu	Ala	Lys	Ala	Val	Val	His	Gly	Asp	Val	Glu	180	185	190	
Gly	Gly	Ser	Val	Met	Ala	Gly	Gln	Ile	Ala	Gly	Leu	Val	Ser	Lys	Glu	195	200	205	
Glu	Thr	Ala	Glu	Glu	Ile	Leu	Lys	Asp	Leu	Tyr	Tyr	Gly	Ala	Ala	Lys	210	215	220	
Lys	Ile	Gln	Glu	Glu	Ala	Ser	Arg	Trp	Thr	Gly	Val	Val	Arg	Asn	Asp	225	230	235	240

(2) INFORMATION FOR SEQ ID NO:170:

(i) SEQUENCE CHARACTERISTICS:

-257-

(A) LENGTH: 243 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

5

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

15

Met Lys Leu Glu His Lys Asn Ile Phe Ile Thr Gly Ser Ser Arg Gly  
 1 5 10 15

Ile Gly Leu Ala Ile Ala His Lys Phe Ala Gln Ala Gly Ala Asn Ile  
 20 25 30

20

Val Leu Asn Ser Arg Gly Ala Ile Ser Glu Glu Leu Leu Ala Glu Phe  
 35 40 45

25

Ser Asn Tyr Gly Ile Lys Val Val Pro Ile Ser Gly Asp Val Ser Asp  
 50 55 60

Phe Ala Asp Ala Lys Arg Met Ile Asp Gln Ala Ile Ala Glu Leu Gly  
 65 70 75 80

30

Ser Val Asp Val Leu Val Asn Asn Ala Gly Ile Thr Gln Asp Thr Leu  
 85 90 95

Met Leu Lys Met Thr Glu Ala Asp Phe Glu Lys Val Leu Lys Val Asn  
 100 105 110

35

Leu Thr Gly Ala Phe Asn Met Thr Gln Ser Val Leu Lys Pro Met Met  
 115 120 125

Lys Ala Arg Glu Gly Ala Ile Ile Asn Met Ser Ser Val Val Gly Leu  
 130 135 140

40

Met Gly Asn Ile Gly Gln Ala Asn Tyr Ala Ala Ser Lys Ala Gly Leu  
 145 150 155 160

45

Ile Gly Phe Thr Lys Ser Val Ala Arg Glu Val Ala Ser Arg Asn Ile  
 165 170 175

Arg Val Asn Val Ile Ala Pro Gly Met Ile Glu Ser Asp Met Thr Ala  
 180 185 190

50

Ile Leu Ser Asp Lys Ile Lys Glu Ala Thr Leu Ala Gln Ile Pro Met  
 195 200 205

Lys Glu Phe Gly Gln Ala Glu Gln Val Ala Asp Leu Thr Val Phe Leu  
 210 215 220

55

Ala Gly Gln Asp Tyr Leu Thr Gly Gln Val Ile Ala Ile Asp Gly Gly  
 225 230 235 240

60

Leu Ser Met

-258-

## (2) INFORMATION FOR SEQ ID NO:171:

5 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 306 amino acids  
     (B) TYPE: amino acid  
     (C) STRANDEDNESS: not relevant  
     (D) TOPOLOGY: not relevant

10 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

20	Met Thr Lys Thr Ala Phe Leu Phe Ala Gly Gln Gly Ala Gln Tyr Leu	1 5 10 15
	Gly Met Gly Arg Asp Phe Tyr Asp Gln Tyr Pro Ile Val Lys Glu Thr	20 25 30
25	Ile Asp Arg Ala Ser Gln Val Leu Gly Tyr Asp Leu Arg Tyr Leu Ile	35 40 45
	Asp Thr Glu Glu Asp Lys Leu Asn Gln Thr Arg Tyr Thr Gln Pro Ala	50 55 60
30	Ile Leu Ala Thr Ser Val Ala Ile Tyr Arg Leu Leu Gln Glu Lys Gly	65 70 75 80
35	Tyr Gln Pro Asp Met Val Ala Gly Leu Ser Leu Gly Glu Tyr Ser Ala	85 90 95
	Leu Val Ala Ser Gly Ala Leu Asp Phe Glu Asp Ala Val Ala Leu Val	100 105 110
40	Ala Lys Arg Gly Ala Tyr Met Glu Glu Ala Ala Pro Ala Asp Ser Gly	115 120 125
	Lys Met Val Ala Val Leu Asn Thr Pro Val Glu Val Ile Glu Glu Ala	130 135 140
45	Cys Gln Lys Ala Ser Glu Leu Gly Val Val Thr Pro Ala Asn Tyr Asn	145 150 155 160
50	Thr Pro Ala Gln Ile Val Ile Ala Gly Glu Val Val Ala Val Asp Arg	165 170 175
	Ala Val Glu Leu Leu Gln Glu Ala Gly Ala Lys Arg Leu Ile Pro Leu	180 185 190
55	Lys Val Ser Gly Pro Phe His Thr Ser Leu Leu Glu Pro Ala Ser Gln	195 200 205
	Lys Leu Ala Glu Thr Leu Ala Gln Val Ser Phe Ser Asp Phe Thr Cys	210 215 220
60	Pro Leu Val Gly Asn Thr Glu Ala Ala Val Met Gln Lys Glu Asp Ile	

-259-

225                      230                      235                      240  
 Ala Gln Leu Leu Thr Arg Gln Val Lys Glu Pro Val Arg Phe Tyr Glu  
                                  245                      250                      255  
 5 Ser Ile Gly Val Met Gln Glu Ala Gly Ile Ser Asn Phe Ile Glu Ile  
                                  260                      265                      270  
 Gly Pro Gly Lys Val Leu Ser Gly Phe Val Lys Lys Ile Asp Gln Thr  
 10                      275                      280                      285  
 Ala His Leu Ala His Val Glu Asp Gln Ala Ser Leu Val Ala Leu Leu  
                                  290                      295                      300  
 15 Glu Lys  
      305

## (2) INFORMATION FOR SEQ ID NO:172:

20 (i) SEQUENCE CHARACTERISTICS:  
      (A) LENGTH: 318 amino acids  
      (B) TYPE: amino acid  
      (C) STRANDEDNESS: not relevant  
      (D) TOPOLOGY: not relevant

25 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

30 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

35 Met Lys Leu Asn Arg Val Val Val Thr Gly Tyr Gly Val Thr Ser Pro  
      1                      5                      10                      15  
 Ile Gly Asn Thr Pro Glu Glu Phe Trp Asn Ser Leu Ala Thr Gly Lys  
                                  20                      25                      30  
 40 Ile Gly Ile Gly Gly Ile Thr Lys Phe Asp His Ser Asp Phe Asp Val  
                                  35                      40                      45  
 45 His Asn Ala Ala Glu Ile Gln Asp Phe Pro Phe Asp Lys Tyr Phe Val  
      50                      55                      60  
 Lys Lys Asp Thr Asn Arg Phe Asp Asn Tyr Ser Leu Tyr Ala Leu Tyr  
      65                      70                      75                      80  
 50 Ala Ala Gln Glu Ala Val Asn His Ala Asn Leu Asp Val Glu Ala Leu  
                                  85                      90                      95  
 Asn Arg Asp Arg Phe Gly Val Ile Val Ala Ser Gly Ile Gly Gly Ile  
                                  100                      105                      110  
 55 Lys Glu Ile Glu Asp Gln Val Leu Arg Leu His Glu Lys Gly Pro Lys  
                                  115                      120                      125  
 60 Arg Val Lys Pro Met Thr Leu Pro Lys Ala Leu Pro Asn Met Ala Ser  
      130                      135                      140



-260-

Gly Asn Val Ala Met Arg Phe Gly Ala Asn Gly Val Cys Lys Ser Ile  
 145 150 155 160  
 Asn Thr Ala Cys Ser Ser Ser Asn Asp Ala Ile Gly Asp Ala Phe Arg  
 165 170 175  
 Ser Ile Lys Phe Gly Phe Gln Asp Val Met Leu Val Gly Gly Thr Glu  
 180 185 190  
 Ala Ser Ile Thr Pro Phe Ala Ile Ala Gly Phe Gln Ala Leu Thr Ala  
 195 200 205  
 Leu Ser Thr Thr Glu Asp Pro Thr Arg Ala Ser Ile Pro Phe Asp Lys  
 210 215 220  
 Asp Arg Asn Gly Phe Val Met Gly Glu Gly Ser Gly Met Leu Val Leu  
 225 230 235 240  
 Glu Ser Leu Glu His Ala Glu Lys Arg Gly Ala Thr Ile Leu Ala Glu  
 245 250 255  
 Val Val Gly Tyr Gly Asn Thr Cys Asp Ala Tyr His Met Thr Ser Pro  
 260 265 270  
 His Pro Glu Gly Gln Gly Ala Ile Lys Ala Ile Lys Leu Ala Leu Glu  
 275 280 285  
 Glu Ala Glu Ile Ser Pro Glu Gln Val Ala Met Leu Met Leu Thr Glu  
 290 295 300  
 Arg Gln Leu Leu Pro Met Lys Lys Glu Lys Val Val Leu Ser  
 305 310 315

## (2) INFORMATION FOR SEQ ID NO:173:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 396 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

Met Gln Ala Val Glu His Phe Ile Lys Gln Phe Val Pro Glu His Tyr  
 1 5 10 15  
 Asp Leu Phe Leu Asp Leu Ser Arg Glu Thr Lys Thr Phe Ser Gly Lys  
 20 25 30  
 Val Thr Ile Thr Gly Gln Ala Gln Ser Asp Arg Ile Ser Leu His Gln  
 35 40 45  
 Lys Asp Leu Glu Ile Thr Ser Val Glu Val Ala Gly Gln Ala Arg Pro  
 50 55 60

-261-

	Phe Thr Val Asp His Asp Asn Glu Ala Leu His Ile Glu Leu Ala Glu	65	70	75	80
5	Ala Gly Gln Val Glu Leu Val Leu Ala Phe Ser Gly Lys Ile Thr Asp	85	90	95	
	Asn Met Thr Gly Ile Tyr Pro Ser Tyr Tyr Thr Val Asp Gly Val Lys	100	105	110	
10	Lys Glu Val Leu Ser Thr Gln Phe Glu Ser His Phe Ala Arg Glu Ala	115	120	125	
	Phe Pro Cys Val Asp Glu Pro Glu Ala Lys Ala Thr Phe Asp Leu Ser	130	135	140	
15	Leu Arg Phe Asp Gln Ala Glu Gly Glu Leu Ala Leu Ser Asn Met Pro	145	150	155	160
	Glu Ile Asp Val Glu Asn Arg Lys Glu Thr Gly Ile Trp Lys Phe Glu	165	170	175	
	Thr Thr Pro Arg Met Ser Ser Tyr Leu Leu Ala Phe Val Ala Gly Asp	180	185	190	
25	Leu Gln Gly Val Thr Ala Lys Thr Lys Asn Gly Thr Leu Val Gly Val	195	200	205	
	Tyr Ser Thr Lys Ala His Pro Leu Ser Asn Leu Asp Phe Ser Leu Asp	210	215	220	
30	Ile Ala Val Arg Ser Ile Glu Phe Tyr Glu Asp Tyr Tyr Gly Val Lys	225	230	235	240
	Tyr Pro Ile Pro Gln Ser Leu His Ile Ala Leu Pro Asp Phe Ser Ala	245	250	255	
	Gly Ala Met Glu Asn Trp Gly Leu Val Thr Tyr Arg Glu Val Tyr Leu	260	265	270	
40	Val Val Asp Glu Asn Ser Thr Phe Ala Ser Arg Gln Gln Val Ala Leu	275	280	285	
	Val Val Ala His Glu Leu Ala His Gln Trp Phe Gly Asn Leu Val Thr	290	295	300	
45	Met Lys Trp Trp Asp Asp Leu Trp Leu Asn Glu Ser Phe Ala Asn Met	305	310	315	320
	Met Glu Tyr Val Cys Val Asp Thr Ile Glu Pro Ser Trp Asn Ile Phe	325	330	335	
	Glu Asp Phe Gln Thr Gly Gly Val Pro Leu Ala Leu Glu Arg Asp Ala	340	345	350	
55	Thr Asp Gly Val Gln Ser Val His Val Glu Val Lys His Pro Asp Glu	355	360	365	
	Ile Asn Thr Leu Phe Asp Gly Ala Ile Val Tyr Ala Arg Lys Arg Leu	370	375	380	

-262-

Met His Met Leu Arg Val Ala Arg Asp Ala Asp Leu  
 385 390 395

## (2) INFORMATION FOR SEQ ID NO:174:

5

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 127 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

10

## (ii) MOLECULE TYPE: peptide

15

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

Met Asp Phe Leu Leu Phe Tyr Asp Ser Lys Lys Lys Gly Asp Thr Met  
 1 5 10 15  
 Thr Tyr Leu Glu Lys Trp Phe Asp Phe Asn Arg Arg Gln Lys Glu Ile  
 20 25 30  
 Glu Ser Leu Leu Glu Glu Thr Ile Ala Gln Gln Ser Glu Gln Ser Leu  
 35 40 45  
 Thr Leu Lys Glu Phe Tyr Leu Leu Tyr Tyr Leu Asp Leu Ala Glu Glu  
 50 55 60  
 Lys Ser Leu Arg Gln Ile Asp Leu Pro Asp Lys Leu His Leu Ser Pro  
 65 70 75 80  
 Ser Ala Val Ser Arg Met Val Ala Arg Leu Glu Ala Lys Asn Cys Gly  
 85 90 95  
 Leu Leu Ser Arg Met Cys Cys His Gln Asp Arg Arg Ser Ser Phe Ile  
 100 105 110  
 Cys Leu Thr Asn Asp Gly Gln Lys Thr Leu Ala Ser Leu Gln Lys  
 115 120 125

45

## (2) INFORMATION FOR SEQ ID NO:175:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 373 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

50

## (ii) MOLECULE TYPE: peptide

55

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

-263-

	Met	Leu	Tyr	Asp	Tyr	Gly	Asn	Ser	Val	Trp	Leu	Ala	Ser	Met	Gly	Thr	
	1				5					10					15		
5	Ile	Gly	Gln	Thr	Val	Leu	Gly	Met	Tyr	Gln	Ile	Ser	Glu	Leu	Val	Thr	
				20					25					30			
	Ser	Ile	Leu	Val	Asn	Pro	Phe	Gly	Gly	Val	Ile	Ser	Asp	Arg	Phe	Ser	
			35					40					45				
10	Arg	Arg	Lys	Ile	Leu	Met	Thr	Ala	Asp	Leu	Val	Cys	Gly	Ile	Leu	Cys	
		50					55					60					
	Leu	Ala	Ile	Ser	Phe	Ile	Arg	Asn	Asp	Ser	Trp	Met	Ile	Gly	Ala	Leu	
15	65					70					75					80	
	Ile	Val	Ala	Asn	Ile	Val	Gln	Ala	Ile	Ala	Phe	Ala	Phe	Ser	Arg	Thr	
				85						90					95		
	Ala	Asn	Lys	Ala	Ile	Ile	Thr	Glu	Val	Val	Glu	Lys	Asn	Glu	Ile	Val	
20				100					105					110			
	Ile	Tyr	Asn	Ser	Arg	Leu	Glu	Leu	Val	Leu	Gln	Val	Val	Gly	Val	Ser	
			115				120						125				
25	Ser	Pro	Val	Leu	Ser	Phe	Leu	Val	Leu	Gln	Phe	Ala	Ser	Leu	His	Met	
		130					135					140					
	Thr	Leu	Leu	Leu	Asp	Ser	Leu	Thr	Phe	Phe	Ile	Ala	Phe	Val	Leu	Val	
30	145					150					155					160	
	Ala	Phe	Leu	Pro	Lys	Glu	Glu	Ala	Lys	Val	Gln	Glu	Lys	Lys	Ala	Phe	
				165						170					175		
	Thr	Gly	Arg	Asp	Ile	Phe	Val	Asp	Ile	Lys	Asp	Gly	Leu	His	Tyr	Ile	
35				180					185					190			
	Trp	His	Gln	Gln	Glu	Ile	Phe	Phe	Leu	Leu	Leu	Val	Ala	Ser	Ser	Val	
			195					200					205				
40	Asn	Phe	Phe	Phe	Ala	Ala	Phe	Glu	Phe	Leu	Leu	Pro	Phe	Ser	Asn	Gln	
		210					215					220					
	Leu	Tyr	Gly	Ser	Glu	Gly	Ala	Tyr	Ala	Ser	Ile	Leu	Thr	Met	Gly	Ala	
45	225					230					235					240	
	Ile	Gly	Ser	Ile	Ile	Gly	Ala	Leu	Leu	Ala	Ser	Lys	Ile	Lys	Ala	Asn	
				245						250					255		
	Ile	Tyr	Asn	Leu	Leu	Ile	Leu	Leu	Ala	Leu	Thr	Gly	Val	Gly	Val	Phe	
50				260					265					270			
	Met	Met	Gly	Leu	Pro	Leu	Pro	Thr	Phe	Leu	Ser	Phe	Ser	Gly	Asn	Leu	
			275					280					285				
55	Val	Cys	Glu	Leu	Phe	Met	Thr	Ile	Phe	Asn	Ile	His	Phe	Phe	Thr	Gln	
		290					295					300					
	Val	Gln	Thr	Lys	Val	Glu	Ser	Glu	Phe	Leu	Gly	Arg	Val	Leu	Ser	Thr	
60	305					310					315					320	
	Ile	Phe	Thr	Leu	Ala	Ile	Leu	Phe	Met	Pro	Ile	Ala	Lys	Gly	Phe	Met	

-264-

[illegible]

(2) INFORMATION FOR SEQ ID NO:176:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 427 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

20 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

	Met 1	Ser	Val	Ser	Phe 5	Glu	Asn	Lys	Glu	Thr 10	Asn	Arg	Gly	Val	Leu 15	Thr
35	Phe	Thr	Ile	Ser 20	Gln	Asp	Gln	Ile	Lys 25	Pro	Glu	Leu	Asp	Arg 30	Val	Phe
	Lys	Ser	Val 35	Lys	Lys	Ser	Leu	Asn 40	Val	Pro	Gly	Phe	Arg 45	Lys	Gly	His
40	Leu	Pro 50	Arg	Pro	Ile	Phe	Asp 55	Gln	Lys	Phe	Gly	Glu 60	Glu	Ala	Leu	Tyr
	Gln 65	Asp	Ala	Met	Asn 70	Ala	Leu	Leu	Pro	Asn 75	Ala	Tyr	Glu	Ala	Ala 80	Val
45	Lys	Glu	Ala	Gly	Leu 85	Glu	Val	Val	Ala	Gln 90	Pro	Lys	Ile	Asp	Val 95	Thr
	Ser	Met	Glu	Lys 100	Gly	Gln	Asp	Trp 105	Val	Ile	Thr	Ala	Glu 110	Val	Val	Thr
50	Lys	Pro	Glu 115	Val	Lys	Leu	Gly	Asp 120	Tyr	Lys	Asn	Leu	Glu 125	Val	Ser	Val
	Asp 130	Val	Glu	Lys	Glu	Val	Thr 135	Asp	Ala	Asp	Val	Glu 140	Glu	Arg	Ile	Glu
55	Arg 145	Glu	Arg	Asn	Asn 150	Leu	Ala	Glu	Leu	Val	Ile 155	Lys	Glu	Ala	Ala 160	Ala
60	Glu	Asn	Gly	Asp	Thr	Val	Val	Ile	Asp	Phe	Val	Gly	Ser	Ile	Asp	Gly

-265-

	165	170	175
	Val Glu Phe Asp Gly Gly Lys Gly Glu Asn Phe Ser Leu Gly Leu Gly		
	180	185	190
5	Ser Gly Gln Phe Ile Pro Gly Phe Glu Asp Gln Leu Val Gly His Ser		
	195	200	205
	Ala Gly Glu Thr Val Asp Val Ile Val Thr Phe Pro Glu Asp Tyr Gln		
10	210	215	220
	Ala Glu Asp Leu Ala Gly Lys Glu Ala Lys Phe Val Thr Thr Ile His		
	225	230	235
15	Glu Val Lys Ala Lys Glu Val Pro Ala Leu Asp Asp Glu Leu Ala Lys		
	245	250	255
	Asp Ile Asp Glu Glu Val Glu Thr Leu Ala Asp Leu Lys Glu Lys Tyr		
20	260	265	270
	Arg Lys Glu Leu Ala Ala Ala Lys Glu Glu Thr Tyr Lys Asp Ala Val		
	275	280	285
	Glu Gly Ala Ala Ile Asp Thr Ala Val Glu Asn Ala Glu Ile Val Glu		
25	290	295	300
	Leu Pro Glu Glu Met Ile His Glu Glu Val His Arg Ser Val Asn Glu		
	305	310	315
30	Phe Leu Gly Asn Leu Gln Arg Gln Gly Ile Asn Pro Asp Met Tyr Phe		
	325	330	335
	Gln Ile Thr Gly Thr Thr Gln Glu Asp Leu His Asn Gln Tyr Gln Ala		
35	340	345	350
	Glu Ala Glu Ser Arg Thr Lys Thr Asn Leu Val Ile Glu Ala Val Ala		
	355	360	365
	Lys Ala Glu Gly Phe Asp Ala Ser Glu Glu Glu Ile Gln Lys Glu Val		
40	370	375	380
	Glu Gln Leu Ala Ala Asp Tyr Asn Met Glu Val Ala Gln Val Gln Asn		
	385	390	395
45	Leu Leu Ser Ala Asp Met Leu Lys His Asp Ile Thr Ile Lys Lys Ala		
	405	410	415
	Val Glu Leu Ile Thr Ser Thr Ala Thr Val Lys		
50	420	425	

## (2) INFORMATION FOR SEQ ID NO:177:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

-266-

(iv) ANTI-SENSE: NO

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

Gly Gly Asp Lys Asp Phe Leu Thr Ser Ile Cys Leu Thr Asn Asp Pro  
 1 5 10 15  
 Phe Leu Gly Phe Arg Ala Leu Arg Ile Ser Ile Ser Glu Thr Gly Asp  
 20 25 30  
 Ala Met Phe Arg Thr Gln Ile Arg Ala Leu Leu Arg Ala Ser Val His  
 35 40 45  
 Gly Gln Leu Arg Ile Met Phe Pro Met Val Ala Leu Leu Lys Glu Phe  
 50 55 60  
 Arg Ala Ala Lys Ala Val Phe Asp Glu Glu Lys Ala Asn Leu Leu Ala  
 65 70 75 80  
 Glu Gly Val Ala Val Ala Asp Asn Ile Gln Val Gly Ile Met Ile Glu  
 85 90 95  
 Ile Pro Ala Ala Ala Met Leu Ala Asp Gln Phe Ala Lys Glu Val Asp  
 100 105 110  
 Phe Phe Ser Ile Gly Thr Asn Asp Leu Ile Gln Tyr Thr Met Ala Ala  
 115 120 125  
 Asp Arg Met Asn Glu Gln Val Ser Tyr Leu Tyr Gln Pro Tyr Asn Pro  
 130 135 140  
 Ser Ile Leu Arg Leu Ile Asn Asn Val Ile Lys Ala Ala His Ala Glu  
 145 150 155 160  
 Gly Lys Trp Ala Gly Met Cys Gly Glu Met Ala Gly Asp Gln Gln Ala  
 165 170 175  
 Val Pro Leu Leu Val Gly Met Gly Leu Asp Glu Phe Ser Met Ser Ala  
 180 185 190  
 Thr Cys Thr Ser Tyr Thr Gln Leu Asp Glu Glu Thr Arg His Ser  
 195 200 205

(2) INFORMATION FOR SEQ ID NO:178:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 283 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

60

-267-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

5	Met	Gln	Met	Ala	Tyr	Arg	Cys	Asn	Leu	Arg	Asn	Asn	Gly	Lys	Arg	Arg	1	5	10	15
	Ile	Gly	Ile	Arg	Glu	Met	Thr	Glu	Met	Leu	Lys	Gly	Ile	Ala	Ala	Ser	20	25	30	
10	Asp	Gly	Val	Ala	Val	Ala	Lys	Ala	Tyr	Leu	Leu	Val	Gln	Pro	Asp	Leu	35	40	45	
	Ser	Phe	Glu	Thr	Ile	Thr	Val	Glu	Asp	Thr	Asn	Ala	Glu	Glu	Ala	Arg	50	55	60	
15	Leu	Asp	Ala	Ala	Leu	Gln	Ala	Ser	Gln	Asp	Glu	Leu	Ser	Val	Ile	Arg	65	70	75	
	Glu	Lys	Ala	Val	Gly	Thr	Leu	Gly	Glu	Glu	Ala	Ala	Gln	Val	Phe	Asp	85	90	95	
20	Ala	His	Leu	Met	Val	Leu	Ala	Asp	Pro	Glu	Met	Ile	Ser	Gln	Ile	Lys	100	105	110	
	Glu	Thr	Ile	Arg	Ala	Lys	Lys	Val	Asn	Ala	Glu	Ala	Gly	Leu	Lys	Glu	115	120	125	
25	Val	Thr	Asp	Met	Phe	Ile	Thr	Ile	Phe	Glu	Gly	Met	Glu	Asp	Asn	Pro	130	135	140	
30	Tyr	Met	Gln	Glu	Arg	Ala	Arg	Asp	Ile	Arg	Asp	Val	Thr	Lys	Arg	Val	145	150	155	
	Leu	Ala	Asn	Leu	Leu	Gly	Lys	Lys	Leu	Pro	Asn	Pro	Ala	Ser	Ile	Asn	165	170	175	
35	Glu	Glu	Val	Ile	Val	Ile	Ala	His	Asp	Leu	Thr	Pro	Ser	Asp	Thr	Ala	180	185	190	
	Gln	Leu	Asp	Lys	Asn	Phe	Val	Lys	Ala	Phe	Val	Thr	Asn	Ile	Gly	Gly	195	200	205	
40	Arg	Thr	Ser	His	Ser	Ala	Ile	Met	Ala	Arg	Thr	Leu	Glu	Ile	Ala	Ala	210	215	220	
45	Val	Leu	Gly	Thr	Asn	Asn	Ile	Thr	Glu	Ile	Val	Lys	Asp	Gly	Asp	Ile	225	230	235	
	Leu	Ala	Val	Asn	Gly	Ile	Thr	Gly	Glu	Val	Ile	Ile	Asn	Pro	Thr	Asp	245	250	255	
50	Glu	Gln	Ala	Ala	Glu	Phe	Lys	Ala	Ala	Gly	Glu	Ala	Tyr	Ala	Thr	Lys	260	265	270	
55	Ala	Glu	Trp	Ala	Leu	Leu	Lys	Asp	Ala	Gln	Gln	275	280							

## (2) INFORMATION FOR SEQ ID NO:179:

60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 168 amino acids



-268-

(B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

5 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO  
 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:  
 15 Met Ile Gly Arg Leu Ala Pro Tyr Asp Lys Gly Gln Ile Ile Tyr Asp  
 1 5 10 15  
 Gly Thr Ser Leu Lys Asp Ile Lys Pro Ser Val Phe Phe Arg Asp Tyr  
 20 25 30  
 Leu Gly Tyr Leu Phe Gln Asp Phe Gly Leu Ile Glu Ser Gln Thr Val  
 35 40 45  
 Lys Glu Asn Leu Asn Leu Gly Leu Val Gly Lys Lys Leu Lys Glu Lys  
 50 55 60  
 25 Glu Lys Ile Ser Leu Met Lys Gln Ala Leu Asn Arg Val Asn Leu Ser  
 65 70 75 80  
 Tyr Leu Asp Leu Lys Gln Pro Ile Phe Glu Leu Ser Gly Gly Glu Ala  
 85 90 95  
 30 Gln Arg Val Ala Leu Ala Lys Ile Ile Leu Lys Asp Pro Pro Leu Ile  
 100 105 110  
 35 Leu Ala Asp Glu Pro Thr Ala Ser Leu Asp Pro Lys Asn Ser Glu Glu  
 115 120 125  
 Leu Leu Ser Ile Leu Glu Ser Leu Lys Asn Pro Asn Arg Thr Ile Ile  
 130 135 140  
 40 Ile Ala Thr His Asn Pro Leu Ile Trp Glu Gln Val Asp Gln Val Ile  
 145 150 155 160  
 45 Arg Val Thr Asp Leu Ser His Arg  
 165

(2) INFORMATION FOR SEQ ID NO:180:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 210 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 55 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 60 (iv) ANTI-SENSE: NO

-269-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

5 Met Lys Ala His Val Ser Tyr Leu Ser Met Gly Glu Lys Arg Phe Val  
 1 5 10 15  
 Tyr Asn Asn Gly Glu Asn Pro Val Ser Thr Gln Tyr Leu Thr Asp Pro  
 20 25 30  
 10 Ile Leu Val Val Phe Thr Pro Thr Ser Thr Gly Asp Ser Phe Ile Ser  
 35 40 45  
 Leu Ser Ser Trp Ser Ile Asn Ala Gly Lys Gln Leu Phe Ile Lys Gly  
 50 55 60  
 15 Tyr Glu Ser Gly Leu Glu Leu Leu Lys Lys Ala Gly Ile Tyr Glu Gln  
 65 70 75 80  
 Val Ser Tyr Leu Lys Glu Gly Arg Ser Val Tyr Leu Thr Arg Tyr Asn  
 85 90 95  
 20 Glu Val Gln Thr Glu Thr Ala Thr Leu Ile Leu Gly Ala Ile Val Gly  
 100 105 110  
 Ile Ala Ser Ser Leu Leu Leu Phe Tyr Ser Val Asn Leu Leu Tyr Phe  
 115 120 125  
 25 Glu Gln Phe Arg Arg Asp Ile Leu Ile Lys Arg Ile Ser Gly Leu Arg  
 130 135 140  
 30 Phe Phe Glu Thr His Ala Gln Tyr Met Val Ser Gln Phe Ala Ser Phe  
 145 150 155 160  
 Val Phe Gly Ala Ser Leu Phe Ile Leu Ser Ser Arg Asp Leu Val Ile  
 165 170 175  
 35 Gly Leu Leu Thr Leu Leu Val Phe Leu Ala Ser Ala Val Leu Thr Leu  
 180 185 190  
 40 Tyr Arg Gln Ala Gln Lys Glu Ser Arg Val Ser Met Thr Ile Met Lys  
 195 200 205  
 Gly Lys  
 210

## 45 (2) INFORMATION FOR SEQ ID NO:181:

## (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 227 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

55 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

-270-

Glu Phe Gln Glu Ala Ser Gln Glu Ser Arg Glu Arg Ser Asp Pro Leu  
 1 5 10 15  
 Asn Ser Tyr Leu Leu Leu Ser Gly Ser Leu Thr Lys Glu Lys Leu Ala  
 5 20 25 30  
 Asp Lys Leu Gly Asp Leu Gly Tyr Lys Ala Ser Ala Asp Arg Lys Ile  
 35 40 45  
 Pro Pro Tyr Phe Leu Ala Phe Arg Ile Leu Leu Asn Pro Leu Ile Leu  
 10 50 55 60  
 Ile Ser Leu Ala Ile Phe Gly Leu Ser Phe Phe Ala Leu Val Ile Ile  
 65 70 75 80  
 Thr Arg Ile Lys Glu Met Arg Ala Ala Gly Ile Lys Leu Phe Ser Gly  
 85 90 95  
 Gln Thr Leu Leu Ser Ile Met Gly His Ser Leu Ser Thr Asp Ile Lys  
 100 105 110  
 Trp Leu Leu Leu Ser Ala Leu Leu Ser Phe Leu Gly Gly Gly Val Val  
 115 120 125  
 Leu Phe Ser Gln Gly Leu Phe Tyr Pro Ile Leu Leu Ala Thr Tyr Gly  
 130 135 140  
 Phe Gly Ile Ser Phe Tyr Leu Leu Phe Leu Leu Ala Ile Ser Ile Leu  
 145 150 155 160  
 Leu Met Leu Leu Tyr Leu Met Ser Leu Asn Lys Ala Leu Val Pro Val  
 165 170 175  
 Ile Arg Gly Arg Phe Pro Leu Leu Met Thr Leu Phe Gln Pro Val Phe  
 180 185 190  
 Ser Val Gly Tyr Ala Lys Thr Gly Leu Thr Ser Tyr Gln Arg Leu Lys  
 195 200 205  
 Glu Leu Glu Ile Ser Gln Trp Gln Asp Arg Val Asp Tyr Tyr His Asp  
 210 215 220  
 Phe Phe Thr  
 225

## (2) INFORMATION FOR SEQ ID NO:182:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 396 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

-271-

	Met	Ser	Lys	Asp	Lys	Lys	Asn	Glu	Asp	Lys	Glu	Thr	Leu	Glu	Glu	Leu	
	1				5					10					15		
5	Lys	Glu	Leu	Ser	Glu	Trp	Gln	Lys	Arg	Asn	Gln	Glu	Tyr	Leu	Lys	Lys	
				20					25					30			
	Lys	Ala	Glu	Glu	Glu	Val	Ala	Leu	Ala	Glu	Glu	Lys	Glu	Lys	Glu	Arg	
			35					40					45				
10	Gln	Ala	Arg	Met	Gly	Glu	Glu	Ser	Glu	Lys	Ser	Glu	Asp	Lys	Gln	Asp	
		50					55					60					
	Gln	Glu	Ser	Glu	Thr	Asp	Gln	Glu	Asp	Ser	Glu	Ser	Ala	Lys	Glu	Glu	
15		65				70					75					80	
	Ser	Glu	Glu	Lys	Val	Ala	Ser	Ser	Glu	Ala	Asp	Lys	Glu	Lys	Glu	Glu	
					85					90				95			
20	Pro	Glu	Ser	Lys	Glu	Lys	Glu	Glu	Gln	Asp	Lys	Lys	Leu	Ala	Lys	Lys	
				100					105					110			
	Ala	Thr	Lys	Glu	Lys	Pro	Ala	Lys	Ala	Lys	Ile	Pro	Gly	Ile	His	Ile	
			115					120					125				
25	Leu	Arg	Ala	Phe	Thr	Ile	Leu	Phe	Pro	Ser	Leu	Leu	Leu	Leu	Ile	Val	
		130					135					140					
	Ser	Ala	Tyr	Leu	Leu	Ser	Pro	Tyr	Ala	Thr	Met	Lys	Asp	Ile	Arg	Val	
30		145				150					155					160	
	Glu	Gly	Thr	Val	Gln	Thr	Thr	Ala	Asp	Asp	Ile	Arg	Gln	Ala	Ser	Gly	
				165					170					175			
35	Ile	Gln	Asp	Ser	Asp	Tyr	Thr	Ile	Asn	Leu	Leu	Leu	Asp	Lys	Ala	Lys	
				180					185					190			
	Tyr	Glu	Lys	Gln	Ile	Lys	Ser	Asn	Tyr	Trp	Val	Glu	Ser	Ala	Gln	Leu	
			195					200					205				
40	Val	Tyr	Gln	Phe	Pro	Thr	Lys	Phe	Thr	Ile	Lys	Val	Lys	Glu	Tyr	Asp	
		210					215					220					
	Ile	Val	Ala	Tyr	Tyr	Ile	Ser	Gly	Glu	Asn	His	Tyr	Pro	Ile	Leu	Ser	
45		225				230					235					240	
	Ser	Gly	Gln	Leu	Glu	Thr	Ser	Ser	Val	Ser	Leu	Asn	Ser	Leu	Pro	Glu	
				245						250				255			
50	Thr	Tyr	Leu	Ser	Val	Leu	Phe	Asn	Asp	Ser	Glu	Gln	Ile	Lys	Val	Phe	
				260					265					270			
	Val	Ser	Glu	Leu	Ala	Gln	Ile	Ser	Pro	Glu	Leu	Lys	Ala	Ala	Ile	Gln	
			275					280					285				
55	Lys	Val	Glu	Leu	Ala	Pro	Ser	Lys	Val	Thr	Ser	Asp	Leu	Ile	Arg	Leu	
		290					295					300					
60	Thr	Met	Asn	Asp	Ser	Asp	Glu	Val	Leu	Val	Pro	Leu	Ser	Glu	Met	Ser	
		305				310					315					320	

-272-

Lys Lys Leu Pro Tyr Tyr Ser Lys Ile Lys Pro Gln Leu Ser Glu Pro  
 325 330 335  
 Ser Val Val Asp Met Glu Ala Gly Ile Tyr Ser Tyr Thr Val Ala Asp  
 340 345 350  
 Lys Leu Ile Met Glu Ala Glu Glu Lys Ala Lys Gln Glu Ala Lys Glu  
 355 360 365  
 Ala Glu Lys Lys Gln Glu Glu Glu Gln Lys Lys Gln Glu Glu Glu Ser  
 370 375 380  
 Asn Arg Asn Gln Thr Asn Gln Arg Ser Ser Arg Arg  
 385 390 395  
 (2) INFORMATION FOR SEQ ID NO:183:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 165 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:  
 Met Glu Arg Val Val Asp Ile Leu Lys Ala Glu Phe Asp Arg Ser Phe  
 1 5 10 15  
 Lys Leu Ile Asn Ser Lys Thr Tyr Pro Val Ser Gly Gly Glu Leu Asn  
 20 25 30  
 Pro Ala Asn Val Asp Ser Glu Ile Glu Ala Phe Ala Gln Leu Gly Val  
 35 40 45  
 Ser Arg Gly Leu Asp Ser Lys Glu Ala His Tyr Leu Ala Asn Leu Tyr  
 50 55 60  
 Gly Ser Asn Ala Pro Lys Val Phe Ala Leu Ala His Ser Leu Glu Gln  
 65 70 75 80  
 Ala Pro Gly Leu Ser Leu Ala Asp Thr Leu Ser Leu His Tyr Ala Met  
 85 90 95  
 Arg Asn Glu Leu Ala Leu Ser Pro Val Asp Phe Leu Leu Arg Arg Thr  
 100 105 110  
 Asn His Met Leu Phe Met Arg Asp Ser Leu Asp Ser Ile Val Glu Pro  
 115 120 125  
 Val Leu Asp Glu Met Gly Arg Phe Tyr Asp Trp Thr Glu Glu Glu Lys  
 130 135 140  
 Ala Thr Tyr Arg Ala Asp Val Glu Ala Ala Leu Ala Asn Asn Asp Leu  
 145 150 155 160

-273-

Ala Glu Leu Lys Asn  
165

5 (2) INFORMATION FOR SEQ ID NO:184:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 233 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

	Met	Asn	Glu	Leu	Phe	Gly	Glu	Phe	Leu	Gly	Thr	Leu	Ile	Leu	Ile	Leu	
	1				5					10					15		
25	Leu	Gly	Asn	Gly	Val	Val	Ala	Gly	Val	Val	Leu	Pro	Lys	Thr	Lys	Ser	
			20					25						30			
	Asn	Ser	Ser	Gly	Trp	Ile	Val	Ile	Thr	Met	Gly	Trp	Gly	Ile	Ala	Val	
			35				40						45				
30	Ala	Val	Ala	Val	Phe	Val	Ser	Gly	Lys	Leu	Ser	Pro	Ala	His	Leu	Asn	
		50					55					60					
	Pro	Ala	Val	Thr	Ile	Gly	Val	Ala	Leu	Lys	Gly	Gly	Leu	Pro	Trp	Ala	
35		65			70						75				80		
	Ser	Val	Leu	Pro	Tyr	Ile	Leu	Ala	Gln	Phe	Ala	Gly	Ala	Met	Leu	Gly	
				85					90					95			
40	Gln	Ile	Leu	Val	Trp	Leu	Gln	Phe	Lys	Pro	His	Tyr	Glu	Ala	Glu	Glu	
			100					105						110			
	Asn	Ala	Gly	Asn	Ile	Leu	Ala	Thr	Phe	Ser	Thr	Gly	Pro	Ala	Ile	Lys	
			115				120					125					
45	Asp	Thr	Val	Ser	Asn	Leu	Ile	Ser	Glu	Ile	Leu	Gly	Thr	Phe	Val	Leu	
		130				135						140					
	Val	Leu	Thr	Ile	Phe	Ala	Leu	Gly	Leu	Tyr	Asp	Phe	Gln	Ala	Gly	Ile	
50		145			150					155					160		
	Gly	Thr	Phe	Ala	Val	Gly	Thr	Leu	Ile	Val	Gly	Ile	Gly	Leu	Ser	Leu	
				165				170						175			
55	Gly	Gly	Thr	Thr	Gly	Tyr	Ala	Leu	Asn	Pro	Ala	Arg	Asp	Leu	Gly	Pro	
				180				185						190			
	Arg	Ile	Met	His	Ser	Ile	Leu	Pro	Ile	Pro	Asn	Lys	Gly	Asp	Gly	Asp	
		195					200						205				
60	Trp	Ser	Tyr	Ala	Trp	Ile	Pro	Val	Val	Gly	Pro	Val	Ile	Gly	Ala	Ala	

-274-

210

215

220

Leu Ala Val Leu Val Leu Ser Leu Phe  
 225 230

5

## (2) INFORMATION FOR SEQ ID NO:185:

## (i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 135 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

## (ii) MOLECULE TYPE: peptide

15

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

Thr Thr Asp Asn Val Ile Asp Leu Phe Glu His Ile Phe Lys Met Phe  
 1 5 10 15  
 Asn Glu Asn Ile Val Met Ala Gly Lys Val Asn Leu Leu Asn Phe Ala  
 20 25 30  
 Asn Leu Ala Ala Tyr Gln Phe Phe Asp Gln Pro Gln Lys Val Ala Leu  
 35 40 45  
 Glu Ile Arg Glu Gly Leu Arg Glu Asp Gln Met Gln Asn Val Arg Val  
 50 55 60  
 Ala Asp Gly Gln Glu Ser Cys Leu Ala Asp Leu Ala Val Ile Ser Ser  
 65 70 75 80  
 Lys Phe Leu Ile Pro Tyr Arg Gly Val Gly Ile Leu Ala Ile Ile Gly  
 85 90 95  
 Pro Val Asn Leu Asp Tyr Gln Gln Leu Ile Asn Gln Ile Asn Val Val  
 100 105 110  
 Asn Arg Val Leu Thr Met Lys Leu Thr Asp Phe Tyr Arg Tyr Leu Ser  
 115 120 125  
 Ser Asn His Tyr Glu Val His  
 130 135

## 50 (2) INFORMATION FOR SEQ ID NO:186:

## (i) SEQUENCE CHARACTERISTICS:

55

- (A) LENGTH: 153 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

## (ii) MOLECULE TYPE: peptide

60

## (iii) HYPOTHETICAL: NO

-275-

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

5 Met Ile Ala Lys Glu Phe Glu Thr Phe Leu Leu Gly Gln Glu Glu Thr  
1 5 10 15

10 Phe Leu Thr Pro Ala Lys Asn Leu Ala Val Leu Ile Asp Thr His Asn  
20 25 30

Ala Asp His Ala Thr Leu Leu Leu Ser Gln Met Thr Tyr Thr Arg Val  
35 40 45

15 Pro Val Val Thr Asp Glu Lys Gln Phe Val Gly Thr Ile Gly Leu Arg  
50 55 60

Asp Ile Met Ala Tyr Gln Met Glu His Asp Leu Ser Gln Glu Ile Met  
65 70 75 80

20 Ala Asp Thr Asp Ile Val His Met Thr Lys Thr Asp Val Ala Val Val  
85 90 95

25 Ser Pro Asp Phe Thr Ile Thr Glu Val Leu His Lys Leu Val Asp Glu  
100 105 110

Ser Phe Leu Pro Val Val Asp Ala Glu Gly Ile Phe Gln Gly Ile Ile  
115 120 125

30 Thr Arg Lys Ser Ile Leu Lys Ala Val Asn Ala Leu Leu His Asp Phe  
130 135 140

Ser Lys Glu Tyr Glu Ile Arg Cys Gln  
145 150

(2) INFORMATION FOR SEQ ID NO:187:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 173 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

45

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

55 Met Ala Lys Gln Thr Ile Ile Val Met Ser Asp Ser His Gly Asp Ser  
1 5 10 15

Leu Ile Val Glu Glu Val Arg Asp Arg Tyr Val Gly Lys Val Asp Ala  
20 25 30

60 Val Phe His Asn Gly Asp Ser Glu Leu Arg Pro Asp Ser Pro Leu Trp  
35 40 45



-276-

Glu Gly Ile Arg Val Val Lys Gly Asn Met Asp Phe Tyr Ala Gly Tyr  
 50 55 60  
 5 Pro Glu Arg Leu Val Thr Glu Leu Gly Ser Thr Lys Ile Ile Gln Thr  
 65 70 75 80  
 His Gly His Leu Phe Asp Ile Asn Phe Asn Phe Gln Lys Leu Asp Tyr  
 85 90 95  
 10 Trp Ala Gln Glu Glu Glu Ala Ala Ile Cys Leu Tyr Gly His Leu His  
 100 105 110  
 Val Pro Ser Ala Trp Met Glu Gly Lys Ile Leu Phe Leu Asn Pro Gly  
 115 120 125  
 15 Ser Ile Ser Gln Pro Arg Gly Thr Ile Arg Glu Cys Leu Tyr Ala Arg  
 130 135 140  
 20 Val Glu Ile Asp Asp Ser Tyr Phe Lys Val Asp Phe Leu Thr Arg Asp  
 145 150 155 160  
 His Glu Val Tyr Pro Gly Leu Ser Lys Glu Phe Ser Arg  
 165 170

25 (2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 189 amino acids  
 (B) TYPE: amino acid  
 30 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

35 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

Met Ser Thr Leu Ala Lys Ile Glu Ala Leu Leu Phe Val Ala Gly Glu  
 1 5 10 15  
 45 Asp Gly Ile Arg Val Arg Gln Leu Ala Glu Leu Leu Ser Leu Pro Pro  
 20 25 30  
 Thr Gly Ile Gln Gln Ser Leu Gly Lys Leu Ala Gln Lys Tyr Glu Lys  
 35 40 45  
 50 Asp Pro Asp Ser Ser Leu Ala Leu Ile Glu Thr Ser Gly Ala Tyr Arg  
 50 55 60  
 55 Leu Val Thr Lys Pro Gln Phe Ala Glu Ile Leu Lys Glu Tyr Ser Lys  
 65 70 75 80  
 Ala Pro Ile Asn Gln Ser Leu Ser Arg Ala Ala Leu Glu Thr Leu Ser  
 85 90 95  
 60 Ile Ile Ala Tyr Lys Gln Pro Ile Thr Arg Ile Glu Ile Asp Ala Ile  
 100 105 110

-277-

Arg Gly Val Asn Ser Ser Gly Ala Leu Ala Lys Leu Gln Ala Phe Asp  
                   115                  120                  125

5 Leu Ile Lys Glu Asp Gly Lys Lys Glu Val Leu Gly Arg Pro Asn Leu  
                   130                  135                  140

Tyr Val Thr Thr Asp Tyr Phe Leu Asp Tyr Met Gly Ile Asn His Leu  
 145                  150                  155                  160

10 Glu Glu Leu Pro Val Ile Asp Glu Leu Glu Ile Gln Ala Gln Glu Ser  
                   165                  170                  175

Gln Leu Phe Gly Glu Arg Ile Glu Glu Asp Glu Asn Gln  
 15                  180                  185

## (2) INFORMATION FOR SEQ ID NO:189:

(i) SEQUENCE CHARACTERISTICS:  
 20 (A) LENGTH: 214 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

25 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

30 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

35 Met Arg Asp Arg Ile Ser Ala Phe Leu Glu Glu Lys Gln Gly Leu Ser  
   1                  5                  10                  15

Val Asn Ser Lys Gln Ser Tyr Lys Tyr Asp Leu Glu Gln Phe Leu Asp  
                   20                  25                  30

40 Met Val Gly Glu Arg Ile Ser Glu Thr Ser Leu Lys Ile Tyr Gln Ala  
                   35                  40                  45

Gln Leu Ala Asn Leu Lys Ile Ser Ala Gln Lys Arg Lys Ile Ser Ala  
 45                  50                  55                  60

Cys Asn Gln Phe Leu Tyr Phe Leu Tyr Gln Lys Gly Glu Val Asp Ser  
 65                  70                  75                  80

50 Phe Tyr Arg Leu Glu Leu Ala Lys Gln Ala Glu Lys Lys Thr Glu Lys  
                   85                  90                  95

Pro Glu Ile Leu Tyr Leu Asp Ser Phe Trp Gln Glu Ser Asp His Pro  
                   100                  105                  110

55 Glu Gly Arg Leu Leu Ala Leu Leu Ile Leu Glu Met Gly Leu Leu Pro  
                   115                  120                  125

Ser Glu Ile Leu Ala Ile Lys Val Ala Asp Ile Asn Leu Asp Phe Gln  
 130                  135                  140

60 Val Leu Arg Ile Ser Lys Ala Ser Gln Gln Arg Ile Val Thr Ile Pro

-278-

145                      150                      155                      160

Thr Ala Leu Leu Ser Glu Leu Glu Pro Leu Met Gly Gln Thr Tyr Leu  
                                  165                      170                      175

5                      Phe Glu Arg Gly Glu Lys Pro Tyr Ser Arg Gln Trp Ala Phe Arg Gln  
                                  180                      185                      190

10                      Leu Glu Ser Phe Val Arg Arg Arg Phe Pro Ser Leu Ser Ala Gln Val  
                                  195                      200                      205

                         Leu Arg Asp Ser Leu Phe  
                                  210

## 15 (2) INFORMATION FOR SEQ ID NO:190:

(i) SEQUENCE CHARACTERISTICS:  
       (A) LENGTH: 239 amino acids  
       (B) TYPE: amino acid  
       (C) STRANDEDNESS: not relevant  
       (D) TOPOLOGY: not relevant

20 (ii) MOLECULE TYPE: peptide

25 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

Met Arg Ile Asn Lys Tyr Ile Ala His Ala Gly Val Ala Ser Arg Arg  
   1                      5                      10                      15

35                      Lys Ala Glu Glu Leu Ile Lys Gln Gly Leu Val Thr Val Asn Gly Gln  
                                  20                      25                      30

40                      Val Val Arg Glu Leu Ala Thr Thr Ile Lys Ser Gly Asp Lys Val Glu  
                                  35                      40                      45

45                      Val Glu Gly Gln Pro Ile Tyr Asn Glu Glu Lys Val Tyr Tyr Leu Leu  
                                  50                      55                      60

                         Asn Lys Pro Arg Gly Val Ile Ser Ser Val Thr Asp Asp Lys Gly Arg  
                                  65                      70                      75                      80

                         Lys Thr Val Val Asp Leu Leu Pro Asn Val Lys Glu Arg Ile Tyr Pro  
                                  85                      90                      95

50                      Val Gly Arg Leu Asp Trp Asp Thr Ser Gly Val Leu Ile Leu Thr Asn  
                                  100                      105                      110

                         Asp Gly Asp Phe Thr Asp Glu Met Ile His Pro Arg Asn Glu Ile Asp  
                                  115                      120                      125

55                      Lys Val Tyr Val Ala Arg Val Lys Gly Val Ala Asn Lys Asp Asn Leu  
                                  130                      135                      140

60                      Arg Pro Leu Thr Arg Gly Leu Glu Ile Asp Gly Lys Lys Thr Lys Pro  
                                  145                      150                      155                      160

-279-

Ala Val Tyr Glu Ile Leu Lys Val Asp Pro Val Lys Asn Arg Ser Val  
165 170 175

5 Val Gln Leu Thr Ile His Glu Gly Arg Asn His Gln Val Lys Lys Met  
180 185 190

Phe Glu Ala Val Gly Leu Gln Val Asp Lys Leu Ser Arg Thr Arg Phe  
195 200 205

10 Gly His Leu Asp Leu Thr Leu Arg Pro Gly Glu Ser Arg Arg Leu Asn  
210 215 220

Lys Lys Glu Ile Ser Gln Leu His Thr Met Ala Val Thr Lys Lys  
225 230 235

15

(2) INFORMATION FOR SEQ ID NO:191:

(i) SEQUENCE CHARACTERISTICS:  
20 (A) LENGTH: 243 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

25 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

Met Asp Ile Lys Leu Lys Arg Phe Leu Lys Asp Pro Gly Leu Ala Leu  
1 5 10 15

35 Cys Ile Trp Phe Leu Ser Thr Lys Met Asp Ile Tyr Asp Val Pro Ile  
20 25 30

40 Thr Glu Val Ile Glu Gln Tyr Leu Ala Tyr Val Ser Thr Leu Gln Ala  
35 40 45

Met Arg Leu Glu Val Thr Gly Glu Tyr Met Val Met Ala Ser Gln Leu  
50 55 60

45 Met Leu Ile Lys Ser Arg Lys Leu Leu Pro Lys Val Ala Glu Val Thr  
65 70 75 80

50 Asp Leu Gly Asp Asp Leu Glu Gln Asp Leu Leu Ser Gln Ile Glu Glu  
85 90 95

Tyr Arg Lys Phe Lys Leu Leu Gly Glu His Leu Glu Ala Lys His Gln  
100 105 110

55 Glu Arg Ala Gln Tyr Tyr Ser Lys Ala Pro Thr Glu Leu Ile Tyr Glu  
115 120 125

Asp Ala Glu Leu Val His Asp Lys Thr Thr Ile Asp Leu Phe Leu Ala  
130 135 140

60 Phe Ser Asn Ile Leu Ala Lys Lys Lys Glu Glu Phe Ala Gln Asn His  
145 150 155 160

-280-

Thr Thr Ile Leu Arg Asp Glu Tyr Lys Ile Glu Asp Met Met Ile Ile  
 165 170 175  
 5 Val Lys Glu Ser Leu Ile Gly Arg Asp Gln Leu Arg Leu Gln Asp Leu  
 180 185 190  
 Phe Lys Glu Ala Gln Asn Val Gln Glu Val Ile Thr Leu Phe Leu Ala  
 195 200 205  
 10 Thr Leu Glu Leu Ile Lys Thr Gln Glu Leu Ile Leu Val Gln Glu Glu  
 210 215 220  
 Ser Phe Gly Asp Ile Tyr Leu Met Glu Lys Lys Glu Glu Ser Gln Val  
 225 230 235 240  
 15 Pro Gln Ser

20 (2) INFORMATION FOR SEQ ID NO:192:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 336 amino acids  
 (B) TYPE: amino acid  
 25 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

30 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

Met Ala Gly Lys Arg Asp Ser Cys Gly Ala Cys Arg Ile Met Thr Asn  
 1 5 10 15  
 40 Lys Ile Tyr Glu Tyr Lys Asp Asp Gln Asn Trp Tyr Val Gly Ser Tyr  
 20 25 30  
 Ser Ile Phe Gly Gly Val Asn Ser Leu Ser Asp Tyr Lys Ala Asp Phe  
 35 40 45  
 45 Pro Leu Phe Glu Phe Ser Lys Ile Phe Gly Asp Glu Glu Tyr Gly Phe  
 50 55 60  
 50 Pro Leu Ser Val Thr Val Leu Arg Tyr Gly Ser Thr Tyr Arg Leu Phe  
 65 70 75 80  
 Ser Phe Val Val Asp Met Leu Asn Gln Glu Met Gly Arg Asn Leu Glu  
 85 90 95  
 55 Val Ile Gln Arg His Gly Ala Leu Leu Leu Val Glu Asn Gly Gln Leu  
 100 105 110  
 Leu Tyr Val Glu Leu Pro Lys Glu Gly Val Asn Val His Asp Phe Phe  
 115 120 125  
 60 Glu Thr Ser Lys Val Arg Glu Thr Leu Leu Ile Ala Thr Arg Asn Glu

-281-

	130		135		140
5	Gly 145	Lys Thr	Lys Glu Phe 150	Arg Ala Ile Phe 155	Asp Lys Leu Gly Tyr Asp 160
	Val	Glu Asn Leu 165	Asn Asp Tyr Pro Asp 170	Leu Pro Glu Val	Ala Glu Thr 175
10	Gly Met Thr 180	Phe Glu Glu Asn Ala 185	Arg Leu Lys Ala	Glu Thr Ile Ser 190	
	Gln Leu Thr 195	Gly Lys Met Val 200	Leu Ala Asp Asp Ser 205	Gly Leu Lys Val	
15	Asp Val Leu 210	Gly Gly Leu Pro Gly Val Trp 215	Ser Ala Arg Phe Ala Gly 220		
	Val Gly Ala Thr 225	Asp Arg Glu Asn Asn Ala 230	Lys Leu Leu His Glu Leu 235		240
20	Ala Met Val 245	Phe Glu Leu Lys Asp Arg 250	Ser Ala Gln Phe His Thr Thr 255		
	Leu Val Val 260	Ala Ser Pro Asn Lys Glu Ser 265	Leu Val Val Glu Ala Asp 270		
	Trp Ser Gly 275	Tyr Ile Asn Phe Glu Pro Lys Gly 280	Glu Asn Gly Phe Gly 285		
30	Tyr Asp Pro 290	Leu Phe Leu Val Gly Glu Thr Gly 295	Glu Ser Ser Ala Glu 300		
	Leu Thr Leu Glu Glu 305	Lys Asn Ser Gln Ser His Arg Ala Leu Ala Val 310			320
35	Lys Lys Leu Leu 325	Glu Val Phe Pro Ser Trp Gln Ser Lys Pro Ser Leu 330			335

40 (2) INFORMATION FOR SEQ ID NO:193:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 219 amino acids  
 (B) TYPE: amino acid  
 45 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

50 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

	Glu 1	Asn 5	Asn 10	Tyr 15	Glu 20	Pro 25	Gln 30	Tyr 35	Ile 40	Asn 45	Ile 50	Arg 55	Gly 60	Lys 65	Gly 70	Pro 75
60	Leu 20	Ile 25	Asn 30	Asp 35	Leu 40	Lys 45	Lys 50	Glu 55	Ala 60	Lys 65	Lys 70	Ala 75	Asn 80	Lys 85	Val 90	Phe 95

-282-

Leu Ala Ser Asp Pro Asp Arg Glu Gly Glu Ala Ile Ser Trp His Leu  
                   35                                  40                                  45  
 5   Ala His Ile Leu Asn Leu Asp Glu Asn Asp Ala Asn Arg Val Val Phe  
                   50                                  55                                  60  
   Asn Glu Ile Thr Lys Asp Ala Val Lys Asn Ala Phe Lys Glu Pro Arg  
   65                                  70                                  75                                  80  
 10   Lys Ile Asp Met Asp Leu Val Asp Ala Gln Gln Ala Arg Arg Ile Leu  
                                   85                                  90                                  95  
   Asp Arg Leu Val Gly Tyr Ser Ile Ser Pro Ile Leu Trp Lys Lys Val  
                                   100                                  105                                  110  
   Lys Lys Gly Leu Ser Ala Gly Arg Val Gln Ser Ile Ala Leu Lys Leu  
                                   115                                  120                                  125  
 20   Ile Ile Asp Arg Glu Asn Glu Ile Asn Ala Phe Gln Pro Glu Glu Tyr  
                                   130                                  135                                  140  
   Trp Thr Val Asp Ala Val Phe Lys Lys Gly Thr Lys Gln Phe His Ala  
   145                                  150                                  155                                  160  
 25   Ser Phe Tyr Gly Val Asp Gly Lys Lys Met Lys Leu Thr Ser Asn Asn  
                                   165                                  170                                  175  
   Glu Val Lys Glu Val Leu Ser Arg Leu Thr Ser Lys Asp Phe Ser Val  
                                   180                                  185                                  190  
   Asp Gln Val Asp Lys Lys Glu Arg Lys Ala Asn Ala Pro Leu Pro Tyr  
                                   195                                  200                                  205  
 35   Thr Thr Ser Ser Met Gln Met Gly Cys Cys Gln  
                                   210                                  215

## (2) INFORMATION FOR SEQ ID NO:194:

40   (i) SEQUENCE CHARACTERISTICS:  
       (A) LENGTH: 236 amino acids  
       (B) TYPE: amino acid  
       (C) STRANDEDNESS: not relevant  
       (D) TOPOLOGY: not relevant  
 45   (ii) MOLECULE TYPE: peptide  
       (iii) HYPOTHETICAL: NO  
 50   (iv) ANTI-SENSE: NO  
       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:  
 55   Met Ser Ile His Ile Ala Ala Gln Gln Gly Glu Ile Ala Asp Lys Ile  
       1                                  5                                  10                                  15  
   Leu Leu Pro Gly Asp Pro Leu Arg Ala Lys Phe Ile Ala Glu Asn Phe  
                                   20                                  25                                  30  
 60   Leu Gly Asp Ala Val Cys Phe Asn Glu Val Arg Asn Met Phe Gly Tyr

-283-

	35	40	45
5	Thr Gly Thr Tyr Lys Gly His Arg Val Ser Val Met Gly Thr Gly Met 50 55 60		
	Gly Met Pro Ser Ile Ser Ile Tyr Ala Arg Glu Leu Ile Val Asp Tyr 65 70 75 80		
10	Gly Val Lys Lys Leu Ile Arg Val Gly Thr Ala Gly Ser Leu Asn Glu 85 90 95		
	Glu Val His Val Arg Glu Leu Val Leu Ala Gln Ala Ala Thr Asn 100 105 110		
15	Ser Asn Ile Val Arg Asn Asp Trp Pro Gln Tyr Asp Phe Pro Gln Ile 115 120 125		
	Ala Ser Phe Asp Leu Leu Asp Lys Ala Tyr His Ile Ala Lys Glu Leu 130 135 140		
20	Gly Met Thr Thr His Val Gly Asn Val Leu Ser Ser Asp Val Phe Tyr 145 150 155 160		
	Ser Asn Tyr Phe Glu Lys Asn Ile Glu Leu Gly Lys Trp Gly Val Lys 165 170 175		
25	Ala Val Glu Met Glu Ala Ala Ala Leu Tyr Tyr Leu Ala Ala Gln Tyr 180 185 190		
	His Val Asp Ala Leu Ala Ile Met Thr Ile Ser Asp Ser Leu Val Asn 195 200 205		
30	Pro Asp Glu Asp Thr Thr Ala Glu Glu Arg Gln Asn Thr Phe Thr Asp 210 215 220		
35	Met Met Lys Val Gly Leu Glu Thr Leu Ile Ala Glu 225 230 235		

## (2) INFORMATION FOR SEQ ID NO:195:

40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 477 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

45 (ii) MOLECULE TYPE: peptide

50 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

55	Ile Ile Phe Pro Ile Leu Thr Gly Thr Tyr Val Ala Arg Val Leu Asp 1 5 10 15
60	Arg Thr Asp Tyr Gly Tyr Phe Asn Ser Val Asp Thr Ile Leu Ser Phe 20 25 30



-284-

	Phe	Leu	Pro	Phe	Ala	Thr	Tyr	Gly	Val	Tyr	Asn	Tyr	Gly	Leu	Arg	Ala
			35					40					45			
5	Ile	Ser	Asn	Val	Lys	Asp	Asn	Lys	Lys	Asp	Leu	Asn	Arg	Thr	Phe	Ser
		50					55					60				
	Ser	Leu	Phe	Tyr	Leu	Cys	Ile	Ala	Cys	Thr	Ile	Leu	Thr	Thr	Ala	Val
	65					70					75					80
10	Tyr	Ile	Leu	Ala	Tyr	Pro	Leu	Phe	Phe	Thr	Asp	Asn	Pro	Ile	Val	Lys
					85					90					95	
	Lys	Val	Tyr	Leu	Val	Met	Gly	Ile	Gln	Leu	Ile	Ala	Gln	Ile	Phe	Ser
15				100				105						110		
	Ile	Glu	Trp	Val	Asn	Glu	Ala	Leu	Glu	Asn	Tyr	Ser	Phe	Leu	Phe	Tyr
		115						120					125			
20	Lys	Thr	Ala	Phe	Ile	Arg	Ile	Leu	Met	Leu	Val	Ser	Ile	Phe	Leu	Phe
		130					135					140				
	Val	Lys	Asn	Glu	His	Asp	Ile	Val	Val	Tyr	Thr	Leu	Val	Met	Ser	Leu
	145					150					155					160
25	Ser	Thr	Leu	Ile	Asn	Tyr	Leu	Ile	Ser	Tyr	Phe	Trp	Ile	Lys	Arg	Asp
					165					170					175	
	Ile	Lys	Leu	Val	Lys	Ile	His	Leu	Ser	Asp	Phe	Lys	Pro	Leu	Phe	Leu
30				180				185						190		
	Pro	Leu	Thr	Ala	Met	Leu	Val	Phe	Ala	Asn	Ala	Asn	Met	Leu	Phe	Thr
			195					200					205			
35	Phe	Leu	Asp	Arg	Leu	Phe	Leu	Val	Lys	Thr	Gly	Ile	Asp	Val	Asn	Val
		210					215					220				
	Ser	Tyr	Tyr	Thr	Ile	Ala	Gln	Arg	Ile	Val	Thr	Val	Ile	Ala	Gly	Val
	225					230					235					240
40	Val	Thr	Gly	Ala	Ile	Gly	Val	Ser	Val	Pro	Arg	Leu	Ser	Tyr	Tyr	Leu
					245					250					255	
	Gly	Lys	Gly	Asp	Lys	Glu	Ala	Tyr	Val	Ser	Leu	Val	Asn	Arg	Gly	Ser
45				260				265						270		
	Arg	Ile	Phe	Asn	Phe	Phe	Ile	Ile	Pro	Leu	Ser	Phe	Gly	Leu	Met	Val
			275					280					285			
50	Leu	Gly	Pro	Asn	Ala	Ile	Leu	Leu	Tyr	Gly	Ser	Glu	Lys	Tyr	Ile	Gly
		290					295					300				
	Gly	Gly	Ile	Leu	Thr	Ser	Leu	Phe	Ala	Phe	Arg	Thr	Ile	Ile	Leu	Ala
	305					310					315					320
55	Leu	Asp	Thr	Ile	Leu	Gly	Ser	Gln	Ile	Leu	Phe	Thr	Asn	Gly	Tyr	Glu
					325					330					335	
	Lys	Arg	Ile	Thr	Val	Tyr	Thr	Val	Phe	Ala	Gly	Leu	Leu	Asn	Leu	Gly
60					340				345					350		
	Leu	Asn	Ser	Leu	Leu	Phe	Phe	Asn	His	Ile	Val	Ala	Pro	Glu	Tyr	Tyr

-285-

355 360 365

Leu Leu Thr Thr Met Leu Ser Glu Thr Ser Leu Leu Val Phe Tyr Ile  
370 375 380

5 Ile Phe Ile His Arg Lys Gln Leu Ile His Leu Gly His Ile Phe Ser  
385 390 395 400

10 Tyr Thr Val Arg Tyr Ser Leu Phe Ser Leu Ser Phe Val Ala Ile Tyr  
405 410 415

Phe Leu Ile Asn Phe Val Tyr Pro Val Asp Met Val Ile Asn Leu Pro  
420 425 430

15 Phe Leu Ile Asn Thr Gly Leu Ile Val Leu Leu Ser Ala Ile Ser Tyr  
435 440 445

Ile Ser Leu Leu Val Phe Thr Lys Asp Ser Ile Phe Tyr Glu Phe Leu  
450 455 460

20 Asn His Val Leu Ala Leu Lys Asn Lys Phe Lys Lys Ser  
465 470 475

## (2) INFORMATION FOR SEQ ID NO:196:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 148 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 30 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

35 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

40 Phe Pro Ile Asp Arg Phe Asp Asp Pro Lys Val Ile Asp Thr Cys Tyr  
1 5 10 15

45 Lys Leu Glu Ser Phe Lys Leu Leu Ser Phe Ser Lys His Lys Asn Ile  
20 25 30

Val Tyr Lys Asp Ser Leu Leu Lys Asp Trp Ile Arg Thr Ala Phe Trp  
35 40 45

50 Leu Leu Leu Arg Pro Val Ser Pro Arg Tyr Phe Ala Asn Lys Ile Glu  
50 55 60

Lys Glu Ile Gln Lys Tyr Ser Arg Glu Asn Gly Gln Tyr Met Ala Phe  
65 70 75 80

55 Ile Pro Ser Lys Phe Lys Glu Lys Glu Val Phe Pro Ser Gly Thr Phe  
85 90 95

60 Asp Lys Thr Ile Asp Leu Pro Phe Glu Asn Leu Ser Leu Pro Ala Pro  
100 105 110

-286-

Glu Lys Phe Asp Thr Ile Leu Thr Gln Phe Tyr Gly Asp Tyr Met Thr  
 115 120 125  
 Leu Pro Pro Glu Glu Lys Arg Phe Tyr Ser His Glu Phe His Ala Tyr  
 130 135 140  
 Lys Leu Glu Asp  
 145

10 (2) INFORMATION FOR SEQ ID NO:197:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 280 amino acids  
 (B) TYPE: amino acid  
 15 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 (ii) MOLECULE TYPE: peptide  
 20 (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO  
 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:  
 Met Asn Phe Thr Leu Ile Asn Trp Arg Ile Arg Met Gln Tyr Leu Glu  
 1 5 10 15  
 30 Lys Lys Glu Ile Lys Glu Ile Gln Leu Ala Leu Leu Asp Tyr Ile Asp  
 20 25 30  
 Glu Thr Cys Lys Lys His Asp Ile Pro Tyr Phe Leu Ser Tyr Gly Thr  
 35 35 40 45  
 Met Leu Gly Ala Ile Arg His Lys Gly Met Ile Pro Trp Asp Asp Asp  
 50 55 60  
 40 Ile Asp Ile Ser Leu Tyr Arg Glu Asp Tyr Glu Arg Leu Leu Lys Ile  
 65 70 75 80  
 Ile Glu Glu Glu Asn His Pro Arg Tyr Lys Val Leu Ser Tyr Asp Thr  
 85 90 95  
 45 Ser Ser Trp Tyr Phe His Asn Phe Ala Ser Ile Leu Asp Thr Ser Thr  
 100 105 110  
 Val Ile Glu Asp His Val Lys Tyr Lys Arg His Asp Thr Ser Leu Phe  
 115 120 125  
 50 Ile Asp Val Phe Pro Ile Asp Arg Phe Thr Asp Leu Ser Ile Val Asp  
 130 135 140  
 Lys Ser Tyr Lys Tyr Val Ala Leu Arg Gln Leu Ala Tyr Ile Lys Lys  
 145 150 155 160  
 Ser Arg Ala Val His Gly Asp Ser Lys Leu Lys Asp Phe Leu Arg Leu  
 165 170 175  
 60 Cys Ser Trp Tyr Ala Leu Arg Phe Val Asn Pro Arg Tyr Phe Tyr Lys  
 180 185 190

-287-

Lys Ile Asp Gln Leu Val Lys Asn Ala Val Thr Asn Thr Pro Gln Tyr  
 195 200 205  
 5 Glu Gly Gly Val Gly Ile Gly Lys Glu Gly Met Lys Glu Ile Phe Pro  
 210 215 220  
 Val Asp Thr Phe Lys Glu Leu Ile Leu Thr Glu Phe Glu Gly Arg Met  
 225 230 235 240  
 10 Leu Pro Val Pro Lys Lys Tyr Asp Gln Phe Leu Thr Gln Met Tyr Gly  
 245 250 255  
 15 Asp Tyr Met Thr Pro Pro Ser Lys Glu Met Gln Glu Trp Tyr Ser His  
 260 265 270  
 Ser Ile Lys Ala Tyr Arg Lys Asn  
 275 280

20 (2) INFORMATION FOR SEQ ID NO:198:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 223 amino acids  
 (B) TYPE: amino acid  
 25 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

30 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

Lys Gly Phe Ile Pro Trp Asp Asp Asp Leu Asp Phe Phe Met Pro Arg  
 1 5 10 15  
 40 Lys Asp Tyr Glu Lys Leu Ala Glu Leu Trp Pro Arg Tyr Ala Asp Glu  
 20 25 30  
 Arg Tyr Phe Leu Ser Lys Ser His Lys Asp Phe Val Asp Arg Asn Leu  
 35 40 45  
 45 Phe Ile Thr Ile Arg Asp Lys Lys Thr Thr Cys Ile Lys Pro Tyr Gln  
 50 55 60  
 50 Gln Asp Leu Asp Leu Pro His Gly Leu Ala Leu Asp Val Leu Pro Leu  
 65 70 75 80  
 Asp Tyr Tyr Pro Lys Asn Pro Ala Glu Arg Lys Lys Gln Val Arg Trp  
 85 90 95  
 55 Ala Leu Ile Tyr Ser Leu Phe Cys Ala Gln Thr Ile Pro Glu Lys His  
 100 105 110  
 Gly Asp Leu Met Lys Trp Gly Ser Arg Ile Leu Leu Gly Leu Thr Pro  
 115 120 125  
 60 Lys Ser Leu Arg Tyr Arg Ile Trp Lys Lys Ala Glu Lys Glu Met Thr

-288-

	130		135		140
	Lys Tyr Asp Leu Ala Asp Cys Asp Gly Ile Thr Glu Leu Cys Ser Gly				
	145		150		155 160
5	Pro Gly Tyr Met Arg Asn Lys Tyr Pro Ile Thr Ser Phe Glu Asp Asn				
		165		170	175
	Leu Phe Leu Pro Phe Glu Gly Thr Glu Met Pro Ile Pro Ile Gly Tyr				
10		180		185	190
	Asp Val Tyr Leu Arg Thr Ala Phe Gly Asp Tyr Met Thr Pro Pro Pro				
		195		200	205
15	Ala Asp Lys Gln Val Pro His His Asp Thr Val Thr Ala Asp Met				
		210		215	220

## (2) INFORMATION FOR SEQ ID NO:199:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 835 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

25 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

30 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

35	Gly Phe Asp Asp Tyr His Pro Ser Cys Gly Arg Ile Leu Ser Val Val
	1 5 10 15
	Thr Ser Gly Gly Glu Asp Ile Ala Asp Ala Ile Ile Ile Leu Ala Val
	20 25 30
40	Val Ile Ile Asn Ala Ala Phe Gly Val Tyr Gln Glu Gly Lys Ala Glu
	35 40 45
45	Glu Ala Ile Glu Ala Leu Lys Ser Met Ser Ser Pro Val Ala Arg Val
	50 55 60
	Leu Arg Asp Gly His Met Ala Glu Ile Asp Ser Lys Glu Leu Val Pro
	65 70 75 80
50	Gly Asp Ile Val Ala Leu Glu Ala Gly Asp Val Val Pro Ala Asp Leu
	85 90 95
	Arg Leu Ile Glu Ala Asn Ser Leu Lys Ile Glu Glu Ala Ala Leu Thr
	100 105 110
55	Gly Glu Ser Val Pro Val Glu Lys Asp Leu Ser Val Asp Leu Ala Thr
	115 120 125
60	Asp Ala Gly Ile Gly Asp Arg Val Asn Met Ala Phe Gln Asn Ser Asn
	130 135 140

-289-

	Val Thr Tyr Gly Arg Gly Met Gly Val Val Val Asn Thr Gly Met Tyr	145	150	155	160
5	Thr Glu Val Gly His Ile Ala Gly Met Leu Gln Asp Ala Asp Glu Thr	165	170	175	
	Asp Thr Pro Leu Lys Gln Asn Leu Asn Asn Leu Ser Lys Val Leu Thr	180	185	190	
10	Tyr Ala Ile Leu Val Ile Ala Leu Val Thr Phe Val Val Gly Val Phe	195	200	205	
	Ile Gln Gly Lys Asn Pro Leu Gly Glu Leu Leu Thr Ser Val Ala Leu	210	215	220	
15	Ala Val Ala Ala Ile Pro Glu Gly Leu Pro Ala Ile Val Thr Ile Val	225	230	235	240
	Leu Ser Leu Gly Thr Gln Val Leu Ala Lys Arg His Ser Ile Val Arg	245	250	255	
20	Lys Leu Pro Ala Val Glu Thr Leu Gly Ser Thr Glu Ile Ile Ala Ser	260	265	270	
	Asp Lys Thr Gly Thr Leu Thr Met Asn Lys Met Thr Val Glu Lys Val	275	280	285	
	Phe Tyr Asp Ala Val Leu His Asp Ser Ala Asp Asp Ile Glu Leu Gly	290	295	300	
30	Leu Glu Met Pro Leu Leu Arg Ser Val Val Leu Ala Asn Asp Thr Lys	305	310	315	320
	Ile Asp Val Glu Gly Asn Leu Ile Gly Asp Pro Thr Glu Thr Ala Phe	325	330	335	
35	Ile Gln Tyr Ala Leu Asp Lys Gly Tyr Asp Val Lys Gly Phe Leu Glu	340	345	350	
	Lys Tyr Pro Arg Val Ala Glu Leu Pro Phe Asp Ser Asp Arg Lys Leu	355	360	365	
	Met Ser Thr Val His Pro Leu Pro Asp Ser Arg Phe Leu Val Ala Val	370	375	380	
45	Lys Gly Ala Pro Asp Gln Leu Leu Lys Arg Cys Leu Leu Arg Asp Lys	385	390	395	400
	Ala Gly Asp Ile Ala Pro Ile Asp Glu Lys Val Thr Asn Leu Ile His	405	410	415	
50	Thr Asn Asn Ser Glu Met Ala His Gln Ala Leu Arg Val Leu Ala Gly	420	425	430	
	Ala Tyr Lys Ile Ile Asp Ser Ile Pro Glu Asn Leu Thr Ser Glu Glu	435	440	445	
	Leu Glu Asn Asp Leu Ile Phe Thr Gly Leu Ile Gly Met Ile Asp Pro	450	455	460	
60	Glu Arg Pro Glu Ala Ala Glu Ala Val Arg Val Ala Lys Glu Ala Gly				

-290-

	465		470		475		480
	Ile Arg Pro Ile Met Ile Thr Gly Asp His Gln Asp Thr Ala Glu Ala						
		485		490			495
5	Ile Ala Lys Arg Leu Gly Ile Ile Asp Ala Asn Asp Thr Glu Gly His						
		500		505			510
10	Val Leu Thr Gly Ala Glu Leu Asn Glu Leu Ser Asp Glu Glu Phe Glu						
		515		520			525
	Lys Val Val Gly Gln Tyr Ser Val Tyr Ala Arg Val Ser Pro Glu His						
		530		535			540
15	Lys Val Arg Ile Val Lys Ala Trp Gln Lys Gln Gly Lys Val Val Ala						
		545		550			555
							560
	Met Thr Gly Asp Gly Val Asn Asp Ala Pro Ala Leu Lys Thr Ala Asp						
		565		570			575
20	Ile Gly Ile Gly Met Gly Ile Thr Gly Thr Glu Val Ser Lys Gly Ala						
		580		585			590
	Ser Asp Met Ile Leu Ala Asp Asp Asn Phe Ala Thr Ile Ile Val Ala						
		595		600			605
25	Val Glu Glu Gly Arg Lys Val Phe Ser Asn Ile Gln Lys Thr Ile Gln						
		610		615			620
30	Tyr Leu Leu Ser Ala Asn Thr Ala Glu Val Leu Thr Ile Phe Leu Ser						
		625		630			635
							640
	Thr Leu Phe Gly Trp Asp Val Leu Gln Pro Val His Leu Leu Trp Ile						
		645		650			655
35	Asn Leu Val Thr Asp Thr Phe Pro Ala Ile Ala Leu Gly Val Glu Pro						
		660		665			670
40	Ala Glu Pro Gly Val Met Asn His Lys Pro Arg Gly Arg Lys Ala Ser						
		675		680			685
	Phe Phe Ser Gly Gly Val Leu Ser Ser Ile Ile Tyr Gln Gly Val Leu						
		690		695			700
45	Gln Ala Ala Leu Val Met Ser Val Tyr Gly Leu Ala Ile Ala Tyr Pro						
		705		710			715
							720
	Val His Val Gly Asp Asn His Ala Ile His Ala Asp Ala Leu Thr Met						
		725		730			735
50	Ala Phe Ala Thr Leu Gly Leu Ile Gln Leu Phe His Ala Tyr Asn Val						
		740		745			750
55	Lys Ser Val Tyr Gln Ser Ile Leu Thr Val Gly Pro Phe Lys Ser Lys						
		755		760			765
	Thr Phe Asn Trp Ser Ile Leu Val Ser Phe Ile Leu Leu Met Ala Thr						
		770		775			780
60	Ile Val Val Glu Pro Leu Glu Gly Ile Phe His Val Thr Lys Leu Asp						
		785		790			795
							800

-291-

Leu Ser Gln Trp Gly Ile Val Met Ala Gly Ser Phe Ser Met Ile Ile  
                     805                    810                    815

5    Ile Val Glu Ile Val Lys Phe Ile Gln Arg Lys Leu Gly Phe Asp Lys  
                     820                    825                    830

Asn Ala Ile  
             835

10    (2) INFORMATION FOR SEQ ID NO:200:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 525 amino acids  
 15    (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide  
 20    (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO

25    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:

Gly Phe Ile Leu Phe Phe Val Leu Leu Gly Ala Val Phe Glu Glu Lys  
 1                      5                      10                      15

30    Met Arg Lys Asn Thr Ser Gln Ala Val Glu Lys Leu Leu Asp Leu Gln  
                     20                      25                      30

35    Ala Lys Thr Ala Glu Val Leu Ser Asp Asp Ser Tyr Val Gln Val Pro  
                     35                      40                      45

Leu Glu Gln Val Lys Val Gly Asp Leu Ile Arg Val Arg Pro Gly Glu  
                     50                      55                      60

40    Lys Ile Ala Val Asp Gly Val Val Val Glu Gly Val Ser Ser Ile Asp  
                     65                      70                      75                      80

Glu Ser Met Val Thr Gly Glu Ser Leu Pro Val Asp Lys Thr Val Gly  
                     85                      90                      95

45    Asp Thr Val Ile Gly Ser Thr Ile Asn His Ser Gly Thr Leu Val Phe  
                     100                      105                      110

50    Arg Ala Glu Lys Val Gly Ser Glu Thr Val Leu Ala Gln Ile Val Asp  
                     115                      120                      125

Phe Val Lys Lys Ala Gln Thr Ser Arg Ala Pro Ile Gln Asp Leu Thr  
                     130                      135                      140

55    Asp Lys Ile Ser Gly Ile Phe Val Pro Val Val Val Ile Leu Gly Ile  
                     145                      150                      155                      160

Met Thr Phe Trp Val Trp Phe Val Leu Leu Arg Asp Ser Val Val Val  
                     165                      170                      175

60    Leu Gly Ala Ser Phe Val Ser Ser Leu Leu Tyr Gly Val Ala Val Leu



-292-

	180	185	190
	Ile Ile Ala Cys Pro Cys Ala Leu Gly Leu Ala Thr Pro Thr Ala Leu		
	195	200	205
5	Met Val Gly Thr Gly Arg Ser Ala Lys Met Gly Val Leu Leu Lys Asn		
	210	215	220
10	Gly Thr Val Leu Gln Glu Ile Gln Lys Val Gln Thr Leu Val Phe Asp		
	225	230	235 240
	Lys Thr Gly Thr Leu Thr Glu Gly Lys Pro Val Val Thr Asp Ile Ile		
	245	250	255
15	Gly Asp Glu Val Glu Val Phe Gly Leu Ala Ala Ser Leu Glu Asp Ala		
	260	265	270
	Ser Gln His Pro Leu Ala Glu Ala Ile Val Lys Arg Ala Ser Glu Ala		
20	275	280	285
	Gly Leu Glu Phe Gln Thr Val Glu Asn Phe Gln Ala Leu His Gly Lys		
	290	295	300
25	Gly Val Ser Gly Arg Ile Asn Gly Lys Gln Val Leu Leu Gly Asn Ala		
	305	310	315 320
	Lys Met Leu Asp Gly Met Asp Ile Ser Asn Thr Tyr Gln Asp Lys Leu		
	325	330	335
30	Glu Glu Leu Glu Lys Glu Ala Lys Thr Val Val Phe Leu Ala Val Asp		
	340	345	350
	Asn Glu Ile Lys Gly Leu Leu Ala Leu Gln Asp Ile Pro Lys Glu Asn		
35	355	360	365
	Ala Lys Leu Ala Ile Ser Gln Leu Lys Lys Arg Gly Leu Arg Thr Val		
	370	375	380
40	Met Leu Thr Gly Asp Asn Ala Gly Val Ala Arg Ala Ile Ala Asp Gln		
	385	390	395 400
	Ile Gly Ile Glu Glu Val Ile Ala Gly Val Leu Pro Glu Glu Lys Ala		
	405	410	415
45	His Glu Ile His Lys Leu Gln Ala Ala Gly Lys Val Ala Phe Val Gly		
	420	425	430
	Asp Gly Ile Asn Asp Ala Pro Ala Leu Ser Val Ala Asp Val Gly Ile		
50	435	440	445
	Ala Met Gly Ala Gly Thr Asp Ile Ala Ile Glu Ser Ala Asp Leu Val		
	450	455	460
55	Leu Thr Thr Asn Asn Leu Leu Gly Val Val Arg Ala Phe Asp Met Ser		
	465	470	475 480
	Lys Lys Thr Phe His Arg Ile Leu Leu Asn Leu Phe Trp Ala Phe Ile		
	485	490	495
60	Tyr Asn Val Val Gly Ile Pro Ile Ala Ala Gly Val Phe Ser Gly Val		
	500	505	510

-293-

Gly Trp Leu Ser Thr Gln Ile Gly Lys Ala Ser Pro Met  
 515 520 525

## 5 (2) INFORMATION FOR SEQ ID NO:201:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 362 amino acids  
 (B) TYPE: amino acid  
 10 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

Asn Asp Ile Ile Glu Phe Met Asp Lys Asn Lys Ile Met Gly Leu Thr  
 1 5 10 15  
 25 Gln Arg Glu Val Lys Glu Arg Gln Ala Glu Gly Leu Val Asn Asp Phe  
 20 25 30  
 Thr Ala Ser Ala Ser Thr Ser Thr Trp Gln Ile Val Lys Arg Asn Val  
 35 40 45  
 30 Phe Thr Leu Phe Asn Ala Leu Asn Phe Ala Ile Ala Leu Ala Leu Ala  
 50 55 60  
 Phe Val Gln Ala Trp Ser Asn Leu Val Phe Phe Ala Val Ile Cys Phe  
 65 70 75 80  
 Asn Ala Phe Ser Gly Ile Val Thr Glu Leu Arg Ala Lys His Met Val  
 85 90 95  
 40 Asp Lys Leu Asn Leu Met Thr Lys Glu Lys Val Lys Thr Ile Arg Asp  
 100 105 110  
 Gly Gln Glu Val Ala Leu Asn Pro Glu Glu Leu Val Leu Gly Asp Val  
 115 120 125  
 45 Ile Arg Leu Ser Ala Gly Glu Gln Ile Pro Ser Asp Ala Leu Val Leu  
 130 135 140  
 50 Glu Gly Phe Ala Glu Val Asn Glu Ala Met Leu Thr Gly Glu Ser Asp  
 145 150 155 160  
 Leu Val Gln Lys Glu Val Asp Gly Leu Leu Leu Ser Gly Ser Phe Leu  
 165 170 175  
 55 Ala Ser Gly Ser Val Leu Ser Gln Val His His Val Gly Ala Asp Asn  
 180 185 190  
 Tyr Ala Ala Lys Leu Met Leu Glu Ala Lys Thr Val Lys Pro Ile Asn  
 195 200 205  
 60 Ser Arg Ile Met Lys Ser Leu Asp Lys Leu Ala Gly Phe Thr Gly Lys

-294-

	210	215	220
	Ile Ile Ile Pro Phe Gly Leu Ala Leu Leu Leu Glu Ala Leu Leu Leu		
	225	230	235 240
5	Lys Gly Leu Pro Leu Lys Ser Ser Val Val Asn Ser Ser Thr Ala Leu		
		245	250 255
	Leu Gly Met Leu Pro Lys Gly Ile Ala Leu Leu Thr Ile Thr Ser Leu		
10		260	265 270
	Leu Thr Ala Val Ile Lys Leu Gly Leu Lys Lys Val Leu Val Gln Glu		
		275	280 285
15	Met Tyr Ser Val Glu Thr Leu Ala Arg Val Asp Met Leu Cys Leu Asp		
		290 295	300
	Lys Thr Gly Thr Ile Thr Gln Gly Lys Met Gln Val Glu Ala Val Leu		
20		305 310	315 320
	Pro Leu Thr Glu Thr Tyr Gly Glu Glu Ala Ile Ala Ser Ile Leu Thr		
		325	330 335
	Ser Tyr Met Ala His Ser Glu Asp Lys Asn Pro Thr Ala Gln Ala Ile		
25		340	345 350
	Arg Gln Arg Leu Trp Glu Met Leu Leu Ile		
		355	360

30 (2) INFORMATION FOR SEQ ID NO:202:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 243 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

35 (ii) MOLECULE TYPE: peptide

40 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

	Ala Ser Asn Ile Met Phe Met Leu Asp Leu Gly Asn His Leu Asp Gln		
	1	5	10 15
50	Trp Ser Leu Lys Lys Thr Ala Thr Asp Leu Glu Gln Ser Leu Leu Ala		
		20	25 30
	Lys Glu Ser Asp Val Phe Leu Val Gln Gly Asp Thr Val Val Ser Ile		
		35	40 45
55	Lys Ser Ser Asp Val Gln Ile Gly Asp Val Leu Ile Leu Ser Gln Gly		
		50	55 60
60	Asn Glu Ile Leu Phe Asp Gly Gln Val Val Ser Gly Leu Gly Met Val		
		65	70 75 80

[illegible]

## 35

—

- 40

## 45

## 50

—

55

60

-296-

Gly Leu Thr Lys Gly Gln Asp Tyr Leu Phe Ile Gly Thr Phe Gln Asp  
 65 70 75 80  
 5 Ser Phe Gln Gly Lys Ile Ile Ser Asn Tyr Val Ser Glu Lys Leu Asn  
 85 90 95  
 Ala Lys Lys Val Val Leu Tyr Thr Asp Asn Ala Ser Asp Tyr Ala Lys  
 100 105 110  
 10 Gly Ile Ala Lys Ser Phe Arg Glu Ser Tyr Lys Gly Glu Ile Val Ala  
 115 120 125  
 Asp Glu Thr Phe Val Ala Gly Asp Thr Asp Phe Gln Ala Ala Leu Thr  
 130 135 140  
 Lys Met Lys Gly Lys Asp Phe Asp Ala Ile Val Val Pro Gly Tyr Tyr  
 145 150 155 160  
 20 Asn Glu Ala Gly Lys Ile Val Asn Gln Ala Arg Gly Met Gly Ile Asp  
 165 170 175  
 Lys Pro Ile Val Gly Gly Asp Gly Phe Asn Gly Glu Glu Phe Val Gln  
 180 185 190  
 25 Gln Ala Thr Ala Glu Lys Ala Ser Asn Ile Tyr Phe Ile Ser Gly Phe  
 195 200 205  
 Ser Thr Thr Val Glu Val Ser Ala Lys Ala Lys Ala Phe Leu Asp Ala  
 210 215 220  
 Tyr Arg Ala Lys Tyr Asn Glu Glu Pro Ser Thr Phe Ala Ala Leu Ala  
 225 230 235 240  
 35 Tyr Asp Ser Val His Leu Val Ala Asn Ala Ala Lys Gly Ala Lys Asn  
 245 250 255  
 Ser Gly Glu Ile Lys Asn Asn Leu Ala Lys Thr Lys Asp Phe Glu Gly  
 260 265 270  
 40 Val Thr Gly Gln Thr Ser Phe Asp Ala Asp His Asn Thr Val Lys Thr  
 275 280 285  
 Ala Tyr Met Met Thr Met Asn Asn Gly Lys Val Glu Ala Ala Glu Val  
 290 295 300  
 Val Lys Pro  
 305

50 (2) INFORMATION FOR SEQ ID NO:204:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 289 amino acids  
 (B) TYPE: amino acid  
 55 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

60 (iii) HYPOTHETICAL: NO

-297-

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

5 Met Leu Gln Gln Leu Val Asn Gly Leu Ile Leu Gly Ser Val Tyr Ala  
1 5 10 15

10 Leu Leu Ala Leu Gly Tyr Thr Met Val Tyr Gly Ile Ile Lys Leu Ile  
20 25 30

Asn Phe Ala His Gly Asp Ile Tyr Met Met Gly Ala Phe Ile Gly Tyr  
35 40 45

15 Phe Leu Ile Asn Ser Phe Gln Met Asn Phe Phe Val Ala Leu Ile Val  
50 55 60

Ala Met Leu Ala Thr Ala Ile Leu Gly Val Val Ile Glu Phe Leu Ala  
65 70 75 80

20 Tyr Arg Pro Leu Arg His Ser Thr Arg Ile Ala Val Leu Ile Thr Ala  
85 90 95

25 Ile Gly Val Ser Phe Leu Leu Glu Tyr Gly Met Val Tyr Leu Val Gly  
100 105 110

Ala Asn Thr Arg Ala Phe Pro Gln Ala Ile Gln Thr Val Arg Tyr Asp  
115 120 125

30 Leu Gly Pro Ile Ser Leu Thr Asn Val Gln Leu Met Ile Leu Gly Ile  
130 135 140

Ser Leu Ile Leu Met Ile Leu Leu Gln Val Ile Val Gln Lys Thr Lys  
145 150 155 160

35 Met Gly Lys Ala Met Arg Ala Val Ser Val Asp Ser Asp Ala Ala Gln  
165 170 175

40 Leu Met Gly Ile Asn Ile Asn Arg Thr Ile Ser Phe Thr Phe Ala Leu  
180 185 190

Gly Ser Ala Leu Ala Gly Ala Ala Gly Val Leu Ile Ala Leu Tyr Tyr  
195 200 205

45 Asn Ser Leu Glu Pro Leu Met Gly Val Thr Pro Gly Leu Lys Ser Phe  
210 215 220

Val Ala Ala Val Leu Gly Gly Ile Gly Ile Ile Pro Gly Ala Ala Leu  
225 230 235 240

50 Gly Gly Phe Val Ile Gly Leu Leu Glu Thr Phe Ala Thr Ala Phe Gly  
245 250 255

55 Met Ser Asp Phe Arg Asp Ala Ile Val Tyr Gly Ile Leu Leu Leu Ile  
260 265 270

Leu Ile Val Arg Pro Ala Gly Ile Leu Gly Lys Asn Val Lys Glu Lys  
275 280 285

60 Val

-298-

## (2) INFORMATION FOR SEQ ID NO:205:

5 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 106 amino acids  
     (B) TYPE: amino acid  
     (C) STRANDEDNESS: not relevant  
     (D) TOPOLOGY: not relevant  
 10 (ii) MOLECULE TYPE: peptide  
     (iii) HYPOTHETICAL: NO  
     (iv) ANTI-SENSE: NO  
 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:  
 20 Ser Gln Asp Gln Thr Trp Tyr Ala Leu Ala Tyr Asp Gly Ala Glu Val  
     1                      5                      10                      15  
     Ile Gly Phe Leu Thr Val Gln Glu Thr Leu Phe Glu Ala Glu Val Leu  
                     20                      25                      30  
 25 Gln Ile Ala Val Lys Gly Ala Tyr Gln Gly Gln Gly Ile Ala Ser Ala  
                     35                      40                      45  
     Leu Phe Ala Gln Leu Pro Thr Asp Lys Glu Ile Phe Leu Glu Val Arg  
                     50                      55                      60  
 30 Gln Ser Asn Gln Arg Ala Gln Ala Phe Tyr Lys Lys Glu Lys Met Ala  
     65                      70                      75                      80  
     Val Ile Ala Glu Arg Lys Ala Tyr Tyr His Asp Pro Val Glu Asp Ala  
                     85                      90                      95  
     Ile Ile Met Lys Arg Glu Ile Asp Glu Gly  
                     100                      105

## 40 (2) INFORMATION FOR SEQ ID NO:206:

45 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 45 amino acids  
     (B) TYPE: amino acid  
     (C) STRANDEDNESS: not relevant  
     (D) TOPOLOGY: not relevant  
     (ii) MOLECULE TYPE: peptide  
 50 (iii) HYPOTHETICAL: NO  
     (iv) ANTI-SENSE: NO  
 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:  
     Lys Thr Leu Lys Gly His Gly Gln Phe Leu His Ala Lys Thr Leu Gly  
     1                      5                      10                      15  
 60 Phe Thr His Pro Arg Thr Gly Lys Thr Leu Glu Phe Lys Ala Asp Ile  
                     20                      25                      30

-299-

Pro Glu Ile Phe Lys Glu Thr Leu Glu Arg Leu Arg Lys  
 35 40 45

5 (2) INFORMATION FOR SEQ ID NO:207:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 163 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

Arg Glu Met Val Val His Pro Ser Ala Gly His Thr Ser Gly Thr Leu  
 1 5 10 15  
 25 Val Asn Ala Leu Met Tyr His Ile Lys Asp Leu Ser Gly Ile Asn Gly  
 20 25 30  
 Val Leu Arg Pro Gly Ile Val His Arg Ile Asp Lys Asp Thr Ser Gly  
 35 40 45  
 30 Leu Leu Met Ile Ala Lys Asn Asp Asp Ala His Leu Val Leu Ala Gln  
 50 55 60  
 Glu Leu Lys Asp Lys Lys Ser Leu Arg Lys Tyr Trp Ala Ile Val His  
 65 70 75 80  
 Gly Asn Leu Pro Asn Asp Arg Gly Val Ile Glu Ala Pro Ile Gly Arg  
 85 90 95  
 40 Ser Glu Lys Asp Arg Lys Lys Gln Ala Val Thr Ala Lys Gly Lys Pro  
 100 105 110  
 Ala Val Thr Arg Phe His Val Leu Glu Arg Phe Gly Asp Tyr Ser Leu  
 115 120 125  
 45 Val Glu Leu Gln Leu Glu Thr Gly Arg Thr His Gln Ile Arg Val His  
 130 135 140  
 Met Ala Tyr Ile Gly His Pro Val Ala Gly Asp Glu Val Tyr Gly Pro  
 50 145 150 155 160  
 Ala Arg Leu

55 (2) INFORMATION FOR SEQ ID NO:208:

(i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 224 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant



-300-

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

10 Leu Gly Thr Arg Gly Ser Ser Arg Val Asp Asn Ile Asn Leu Gln Val  
 1 5 10 15  
 15 Asp Glu Arg Asp Arg Ile Ala Leu Val Gly Lys Asn Gly Ala Gly Lys  
 20 25 30  
 Ser Thr Leu Leu Lys Ile Leu Val Gly Glu Glu Glu Pro Thr Ser Gly  
 35 40 45  
 20 Glu Ile Asn Lys Lys Lys Asp Ile Ser Leu Ser Tyr Leu Ala Gln Asp  
 50 55 60  
 Ser Arg Phe Glu Ser Glu Asn Thr Ile Tyr Asp Glu Met Leu His Val  
 65 70 75 80  
 25 Phe Asn Asp Leu Arg Arg Thr Glu Arg Gln Leu Arg Gln Met Glu Leu  
 85 90 95  
 30 Glu Met Gly Glu Lys Ser Gly Glu Asp Leu Asp Lys Leu Met Ser Asp  
 100 105 110  
 Tyr Asp Arg Leu Ser Glu Asn Phe Arg Gln Ala Gly Gly Phe Thr Tyr  
 115 120 125  
 35 Glu Ala Asp Ile Arg Ala Ile Leu Asn Gly Phe Lys Phe Asp Glu Ser  
 130 135 140  
 Met Trp Gln Met Lys Ile Ala Glu Leu Ser Gly Gly Gln Asn Thr Arg  
 145 150 155 160  
 40 Leu Ala Leu Ala Lys Met Leu Leu Glu Lys Pro Asn Leu Leu Val Leu  
 165 170 175  
 Asp Glu Pro Thr Asn His Leu Asp Ile Glu Thr Ile Ala Trp Leu Glu  
 180 185 190  
 Asn Tyr Leu Val Asn Tyr Ser Gly Ala Leu Ile Ile Val Ser His Asp  
 195 200 205  
 50 Arg Tyr Phe Leu Asp Lys Val Ala Thr Ile Thr Leu Asp Leu Thr Ser  
 210 215 220

(2) INFORMATION FOR SEQ ID NO:209:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 192 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

-301-

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

Ser Thr Thr His His Leu Leu Val Lys Lys Val Asn Gly Leu Leu Val  
 1 5 10 15  
 Arg Trp Lys Asn Ala Cys Arg Gln Asn Cys Lys Gln Thr Phe Xaa Phe  
 20 25 30  
 Val Leu Thr Gln Leu Ile His Ala Asp Lys Trp Thr Val Ser Gly Arg  
 35 40 45  
 Gly Glu Leu His Leu Ser Ile Leu Ile Glu Thr Met Arg Arg Glu Gly  
 50 55 60  
 Tyr Glu Leu Gln Val Ser Arg Pro Glu Val Ile Val Lys Glu Ile Asp  
 65 70 75 80  
 Gly Val Lys Cys Glu Pro Phe Glu Arg Val Gln Ile Asp Thr Pro Glu  
 85 90 95  
 Glu Tyr Gln Gly Ser Val Ile Gln Ser Leu Ser Glu Arg Lys Gly Glu  
 100 105 110  
 Met Leu Asp Met Ile Ser Thr Gly Asn Gly Gln Thr Arg Leu Val Phe  
 115 120 125  
 Leu Val Pro Ala Arg Gly Leu Xaa Trp Ile Leu Asn Val Leu Val Asn  
 130 135 140  
 Asp Ser Trp Leu Arg Tyr His Glu Pro Tyr Leu Arg Pro Ile Leu Ala  
 145 150 155 160  
 Ile Asp Ser Arg Gly Asn Trp Trp Thr Ser Pro Trp Cys Pro Cys Phe  
 165 170 175  
 Tyr Arg Cys Trp Gly Tyr Asn Leu Leu Asn Leu Leu Leu Ser Thr Leu  
 180 185 190

50 (2) INFORMATION FOR SEQ ID NO:210:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 252 amino acids

(B) TYPE: amino acid

55 (C) STRANDEDNESS: not relevant

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

60 (iv) ANTI-SENSE: NO

-302-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:210:

5 Met Phe Gly Phe Phe Lys Lys Asp Lys Ala Val Glu Val Glu Val Pro  
 1 5 10 15  
 Thr Gln Val Pro Ala His Ile Gly Ile Ile Met Asp Gly Asn Gly Arg  
 20 25 30  
 10 Trp Ala Lys Lys Arg Met Gln Pro Arg Val Phe Gly His Lys Ala Gly  
 35 40 45  
 15 Met Glu Ala Leu Gln Thr Val Thr Lys Ala Ala Asn Lys Leu Gly Val  
 50 55 60  
 Lys Val Ile Thr Val Tyr Ala Phe Ser Thr Glu Asn Trp Thr Arg Pro  
 65 70 75 80  
 20 Asp Gln Glu Val Lys Phe Xaa Met Asn Leu Pro Val Glu Phe Tyr Asp  
 85 90 95  
 Asn Tyr Val Pro Glu Leu His Ala Asn Asn Val Lys Ile Gln Met Ile  
 100 105 110  
 25 Gly Glu Thr Asp Arg Leu Pro Lys Gln Thr Phe Glu Ala Leu Thr Lys  
 115 120 125  
 Ala Glu Glu Leu Thr Lys Asn Asn Thr Gly Leu Ile Leu Asn Phe Ala  
 130 135 140  
 30 Leu Asn Tyr Gly Gly Arg Ala Glu Ile Thr Gln Ala Leu Lys Leu Ile  
 145 150 155 160  
 35 Ser Gln Asp Val Leu Asp Ala Lys Ile Asn Pro Gly Asp Ile Thr Glu  
 165 170 175  
 Glu Leu Ile Gly Asn Tyr Leu Phe Thr Gln His Leu Pro Lys Asp Leu  
 180 185 190  
 40 Arg Asp Pro Asp Leu Ile Ile Arg Thr Ser Gly Glu Leu Arg Leu Ser  
 195 200 205  
 45 Asn Phe Leu Pro Trp Gln Gly Ala Tyr Ser Glu Leu Tyr Phe Thr Asp  
 210 215 220  
 Thr Leu Trp Pro Asp Phe Asp Glu Ala Ala Leu Gln Glu Ala Ile Leu  
 225 230 235 240  
 50 Ala Tyr Asn Arg Arg His Arg Arg Phe Gly Gly Val  
 245 250

## (2) INFORMATION FOR SEQ ID NO:211:

- 55 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 67 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 60 (ii) MOLECULE TYPE: peptide

-303-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:211:

10 Val Glu Gln Lys Leu Arg Gly Arg Asn Glu Asn Glu Ile Gln Ser Glu  
 1 5 10 15  
 Asp Ile Gly Ser Leu Val Met Glu Glu Leu Ala Glu Leu Asp Glu Ile  
 20 25 30  
 15 Thr Tyr Val Arg Phe Ala Ser Val Tyr Arg Ser Phe Lys Asp Val Ser  
 35 40 45  
 Glu Leu Glu Ser Leu Leu Gln Gln Ile Thr Gln Ser Ser Lys Lys Lys  
 20 50 55 60  
 Lys Glu Arg  
 65

(2) INFORMATION FOR SEQ ID NO:212:

25

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 75 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

30

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

35

(iv) ANTI-SENSE: NO

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:

45 Val Asp Ser Arg Gln Ala Glu Glu Gly Asn Thr Ile Arg Arg Arg Arg  
 1 5 10 15  
 Glu Cys Asp Glu Cys Gln His Arg Phe Thr Thr Tyr Glu Arg Val Glu  
 20 25 30  
 50 Glu Arg Thr Leu Val Val Val Lys Lys Asp Gly Thr Arg Glu Gln Phe  
 35 40 45  
 Ser Arg Asp Lys Ile Phe Asn Gly Ile Ile Arg Ser Ala Gln Lys Arg  
 50 55 60  
 55 Pro Val Ser Ser Asp Glu Ile Asn Met Val Ile  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:213:

60

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 63 amino acids

-304-

(B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

5 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 10 (iv) ANTI-SENSE: NO  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:  
 15 Phe Ala Gln Val Pro Lys Val Ala Gln Lys Val Met Lys Val Thr Lys  
 1 5 10 15  
 Ala Ala Gly Met Asn Ile Ile Ser Asn Cys Glu Glu Val Ala Gly Gln  
 20 20 25 30  
 Thr Val Phe His Thr His Val His Leu Val Pro Arg Tyr Ser Ala Asp  
 35 40 45  
 Asp Asp Leu Lys Ile Asp Phe Ile Ala His Glu Thr Asp Phe Asp  
 25 50 55 60

## (2) INFORMATION FOR SEQ ID NO:214:

(i) SEQUENCE CHARACTERISTICS:  
 30 (A) LENGTH: 67 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 35 (ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTI-SENSE: NO  
 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:  
 45 Met Ser Asp Cys Ile Phe Cys Lys Ile Ile Ala Gly Glu Ile Pro Ala  
 1 5 10 15  
 Ser Lys Val Tyr Glu Asp Glu Gln Val Leu Ala Phe Leu Asp Ile Ser  
 20 25 30  
 Gln Val Thr Leu Gly His Thr Leu Val Val Pro Lys Glu His Tyr Arg  
 50 35 40 45  
 Asn Leu Leu Glu Met Asp Ala Thr Ser Ala Thr Asn Ser Leu Pro Lys  
 50 55 60  
 55 Tyr Gln Lys  
 65

## (2) INFORMATION FOR SEQ ID NO:215:

60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 212 amino acids

-305-

(B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

5 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

10 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:

15 Ile Gln Ala Val Arg Asp Val Ser Phe Glu Val Asn Glu Gly Glu Val  
 1 5 10 15  
 Val Ser Leu Ile Gly Ala Asn Gly Ala Gly Lys Thr Thr Ile Leu Arg  
 20 20 25 30  
 Thr Leu Ser Gly Leu Val Arg Pro Ser Ser Gly Lys Ile Glu Phe Leu  
 35 40 45  
 Gly Gln Glu Ile Gln Lys Met Pro Ala Gln Lys Ile Val Ala Gly Gly  
 50 55 60  
 25 Leu Ser Gln Val Pro Glu Gly Arg His Val Phe Pro Gly Leu Thr Val  
 65 70 75 80  
 Met Glu Asn Leu Glu Met Gly Ala Phe Leu Lys Lys Asn Arg Glu Glu  
 85 90 95  
 30 Asn Gln Ala Asn Leu Lys Lys Val Phe Ser Arg Phe Pro Arg Leu Glu  
 100 105 110  
 Glu Arg Lys Asn Gln Asp Ala Ala Thr Leu Ser Gly Gly Glu Gln Gln  
 115 120 125  
 Met Leu Ala Met Gly Arg Ala Leu Met Ser Thr Pro Lys Leu Leu Leu  
 130 135 140  
 40 Leu Asp Glu Pro Ser Met Gly Leu Ala Pro Ile Phe Ile Gln Glu Ile  
 145 150 155 160  
 Phe Asp Ile Ile Gln Asp Ile Gln Lys Gln Gly Thr Thr Val Leu Leu  
 165 170 175  
 45 Ile Glu Gln Asn Ala Asn Lys Ala Leu Ala Ile Ser Asp Arg Gly Tyr  
 180 185 190  
 Val Leu Glu Gln Gly Asn Arg Leu Ser Gly Thr Gly Lys Asp Ser Leu  
 195 200 205  
 Ile Arg Gly Val  
 210

55 (2) INFORMATION FOR SEQ ID NO:216:

(i) SEQUENCE CHARACTERISTICS:

60 (A) LENGTH: 43 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant

-306-

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:

Leu Leu Ser Leu Ile Asp Ile Leu Val Asp Gly Arg Tyr Asp Arg Thr  
 1 5 10 15

15 Lys Arg Asn Leu Met Leu Gln Phe Arg Gly Ser Ser Asn Gln Arg Ile  
 20 25 30

Ile Asp Ser Arg Gly Ser Pro Gly Thr Glu Leu  
 35 40

20

(2) INFORMATION FOR SEQ ID NO:217:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 130 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

30

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:

Met Asn Asn Pro Lys Pro Gln Glu Trp Lys Ser Glu Glu Leu Ser Gln  
 1 5 10 15

40

Gly Arg Ile Ile Asp Tyr Lys Ala Phe Asn Phe Val Asp Gly Glu Gly  
 20 25 30

45

Val Arg Asn Ser Leu Tyr Val Ser Gly Cys Met Phe His Cys Glu Gly  
 35 40 45

Cys Tyr Asn Val Ala Thr Trp Ser Phe Asn Ala Gly Ile Pro Tyr Thr  
 50 55 60

50

Ala Glu Leu Glu Glu Gln Ile Met Ala Asp Leu Ala Gln Pro Tyr Val  
 65 70 75 80

Gln Gly Leu Thr Leu Leu Gly Gly Glu Pro Phe Leu Asn Thr Gly Ile  
 85 90 95

55

Leu Leu Pro Leu Val Lys Arg Ile Arg Lys Glu Leu Pro Asp Lys Asp  
 100 105 110

60

Ile Trp Ser Trp Thr Gly Tyr Thr Trp Glu Glu Met Ile Pro Gly Asn  
 115 120 125

-307-

Ser Arg  
130

5 (2) INFORMATION FOR SEQ ID NO:218:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 126 amino acids  
 (B) TYPE: amino acid  
 10 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:

Met Val Asn His Phe Arg Ile Asp Arg Val Gly Met Glu Ile Lys Arg  
 1 5 10 15

25 Glu Val Asn Glu Ile Leu Gln Lys Lys Val Arg Asp Pro Arg Val Gln  
 20 25 30

Gly Val Thr Ile Thr Asp Val Gln Met Leu Gly Asp Leu Ser Val Ala  
 35 40 45

30 Lys Val Tyr Tyr Thr Ile Leu Ser Asn Leu Ala Ser Asp Asn Gln Lys  
 50 55 60

35 Ala Gln Ile Gly Leu Glu Lys Ala Thr Gly Thr Ile Lys Arg Glu Leu  
 65 70 75 80

Gly Arg Asn Leu Lys Leu Tyr Xaa Ile Pro Asp Leu Thr Phe Val Lys  
 85 90 95

40 Asp Glu Ser Ile Glu Xaa Gly Thr Lys Ile Asp Glu Met Leu Arg Asn  
 100 105 110

Leu Asp Lys Thr Lys Glu Glu Gly Val Ala Pro Leu Phe Trp  
 115 120 125

45 (2) INFORMATION FOR SEQ ID NO:219:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 267 amino acids  
 (B) TYPE: amino acid  
 50 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

55 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:



-308-

Phe His His Val Thr Val Leu Leu His Glu Thr Ile Asp Met Leu Asp  
 1 5 10 15  
 Val Lys Pro Glu Gly Ile Tyr Val Asp Ala Thr Leu Gly Gly Ala Gly  
 5 20 25 30  
 His Ser Glu Tyr Leu Leu Ser Lys Leu Ser Glu Lys Gly His Leu Tyr  
 35 40 45  
 Ala Phe Asp Gln Asp Gln Asn Ala Ile Asp Asn Ala Gln Lys Arg Leu  
 10 50 55 60  
 Ala Pro Tyr Ile Glu Lys Gly Met Val Thr Phe Ile Lys Asp Asn Phe  
 15 65 70 75 80  
 Arg His Leu Gln Ala Arg Leu Arg Glu Ala Gly Val Gln Glu Ile Asp  
 85 90 95  
 Gly Ile Cys Tyr Asp Leu Gly Val Ser Ser Pro Gln Leu Asp Gln Arg  
 20 100 105 110  
 Glu Arg Gly Phe Ser Tyr Lys Lys Asp Ala Pro Leu Asp Met Arg Met  
 115 120 125  
 Asn Gln Asp Ala Ser Leu Thr Ala Tyr Glu Val Val Asn His Tyr Asp  
 130 135 140  
 Tyr His Asp Leu Val Arg Ile Phe Phe Lys Tyr Gly Glu Asp Lys Phe  
 145 150 155 160  
 Ser Lys Gln Ile Ala Arg Lys Ile Glu Gln Ala Arg Glu Val Lys Pro  
 165 170 175  
 Ile Glu Thr Thr Thr Glu Leu Ala Glu Ile Ile Lys Leu Val Lys Pro  
 180 185 190  
 Ala Lys Glu Leu Lys Lys Lys Gly His Pro Ala Lys Gln Ile Phe Gln  
 195 200 205  
 Ala Ile Arg Ile Glu Val Asn Asp Glu Leu Gly Ala Ala Asp Glu Ser  
 210 215 220  
 Ile Gln Gln Ala Met Asp Met Leu Ala Leu Asp Gly Arg Ile Ser Val  
 225 230 235 240  
 Ile Thr Phe His Ser Leu Glu Asp Arg Leu Thr Lys Gln Leu Phe Lys  
 245 250 255  
 Xaa Ala Ser Thr Val Glu Val Pro Lys Gly Leu  
 260 265

## (2) INFORMATION FOR SEQ ID NO:220:

- (i) SEQUENCE CHARACTERISTICS:  
 55 (A) LENGTH: 165 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant  
 60 (ii) MOLECULE TYPE: peptide

-309-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

10 Leu Met His Val Thr Val Gly Glu Leu Ile Gly Asn Phe Ile Leu Ile  
 1 5 10 15  
 Thr Gly Ser Phe Ile Leu Leu Leu Val Leu Ile Lys Lys Phe Ala Trp  
 20 25 30  
 15 Ser Asn Ile Thr Gly Ile Phe Glu Glu Arg Ala Glu Lys Ile Ala Ser  
 35 40 45  
 Asp Ile Asp Arg Ala Glu Glu Ala Arg Gln Lys Ala Glu Val Leu Ala  
 50 55 60  
 20 Gln Lys Arg Glu Asp Glu Leu Ala Gly Ser Arg Lys Glu Ala Lys Thr  
 65 70 75 80  
 Ile Ile Glu Asn Ala Lys Glu Thr Ala Glu Gln Ser Lys Ala Asn Ile  
 85 90 95  
 25 Leu Ala Asp Ala Lys Leu Glu Ala Gly His Leu Lys Glu Lys Ala Asn  
 100 105 110  
 30 Gln Glu Ile Ala Gln Asn Lys Val Glu Ala Leu Gln Ser Val Lys Gly  
 115 120 125  
 Glu Val Ala Asp Leu Thr Ile Ser Leu Ala Gly Lys Ile Ile Ser Gln  
 130 135 140  
 35 Asn Leu Asp Ser His Ala His Lys Ala Leu Ile Asp Gln Tyr Ile Asp  
 145 150 155 160  
 Gln Leu Gly Glu Ala  
 165

40

(2) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 629 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

50 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

60 Met Gln Arg Leu Val Ser Leu Leu Ile Trp Ser Leu Leu Glu Thr Ser  
 1 5 10 15  
 Ile Leu Ser Ile His Gly Leu Gly Pro Leu Thr Lys Arg Phe Gly Val

-310-

	20	25	30
5	Ala Leu Glu His His His Met	Ala Asn Tyr Asp	Ala Glu Ala Thr Gly
	35	40	45
	Arg Leu Leu Phe Ile Phe Ile Lys Glu Val Ala Glu Lys His Gly Val		
	50	55	60
10	Thr Asp Leu Ala Arg Leu Asn Ile Asp Leu Ile Ser Pro Asp Ser Tyr		
	65	70	75
	Lys Lys Ala Arg Ile Lys His Ala Thr Ile Tyr Val Lys Asn Gln Val		
	85	90	95
15	Gly Leu Lys Asn Ile Phe Lys Leu Val Ser Leu Ser Asn Thr Lys Tyr		
	100	105	110
	Phe Glu Gly Val Ser Arg Ile Pro Arg Thr Val Leu Asp Ala His Arg		
	115	120	125
20	Glu Gly Leu Ile Leu Gly Ser Ala Cys Ser Glu Gly Glu Val Phe Asp		
	130	135	140
	Val Val Val Ser Gln Gly Val Asp Ala Ala Val Glu Val Ala Lys Tyr		
	145	150	155
25	Tyr Asp Phe Ile Glu Val Met Pro Pro Ala Ile Tyr Ala Pro Leu Ile		
	165	170	175
30	Ala Lys Glu Gln Val Lys Asp Met Glu Glu Leu Gln Thr Ile Ile Lys		
	180	185	190
	Ser Leu Ile Glu Val Gly Asp Arg Leu Gly Lys Pro Val Leu Ala Thr		
	195	200	205
35	Gly Asn Val His Tyr Ile Glu Pro Glu Glu Glu Ile Tyr Arg Glu Ile		
	210	215	220
	Ile Val Arg Ser Leu Gly Gln Gly Ala Met Ile Asn Arg Thr Ile Gly		
	225	230	235
40	His Gly Glu His Ala Gln Pro Ala Pro Leu Pro Lys Ala His Phe Arg		
	245	250	255
	Thr Thr Asn Glu Met Leu Asp Glu Phe Ala Phe Leu Gly Glu Glu Leu		
	260	265	270
	Ala Arg Lys Leu Val Ile Glu Asn Thr Asn Ala Leu Ala Glu Ile Phe		
	275	280	285
50	Glu Pro Val Glu Val Val Lys Gly Asp Leu Tyr Thr Pro Phe Ile Asp		
	290	295	300
	Lys Ala Glu Glu Thr Val Ala Glu Leu Thr Tyr Lys Lys Ala Phe Glu		
	305	310	315
55	Ile Tyr Gly Asn Pro Leu Pro Asp Ile Val Asp Leu Arg Ile Glu Lys		
	325	330	335
60	Glu Leu Thr Ser Ile Leu Gly Asn Gly Phe Ala Val Ile Tyr Leu Ala		
	340	345	350

-311-

Ser Gln Met Leu Val Gln Arg Ser Asn Glu Arg Gly Tyr Leu Val Gly  
 355 360 365  
 5 Ser Arg Gly Ser Val Gly Ser Ser Phe Val Ala Thr Met Ile Gly Ile  
 370 375 380  
 Thr Glu Val Asn Pro Leu Ser Pro His Tyr Val Cys Gly Gln Cys Gln  
 385 390 395 400  
 10 Tyr Ser Glu Phe Ile Thr Asp Gly Ser Tyr Gly Ser Gly Phe Asp Met  
 405 410 415  
 Pro His Lys Asp Cys Pro Asn Cys Gly His Lys Leu Ser Lys Asn Gly  
 420 425 430  
 Gln Asp Ile Pro Phe Glu Thr Phe Leu Gly Phe Asp Gly Asp Lys Val  
 435 440 445  
 20 Pro Asp Ile Asp Leu Asn Phe Ser Gly Glu Asp Gln Pro Ser Ala His  
 450 455 460  
 Leu Asp Val Arg Asp Ile Phe Gly Glu Glu Tyr Ala Phe Arg Ala Gly  
 465 470 475 480  
 25 Thr Val Gly Thr Val Ala Ala Lys Thr Ala Tyr Gly Phe Val Lys Gly  
 485 490 495  
 Tyr Glu Arg Asp Tyr Gly Lys Phe Tyr Arg Asp Ala Glu Val Glu Arg  
 500 505 510  
 30 Leu Ala Gln Gly Ala Ala Gly Val Lys Arg Thr Thr Gly Gln His Pro  
 515 520 525  
 Gly Gly Ile Val Val Ile Pro Asn Tyr Met Asp Val Tyr Asp Phe Thr  
 530 535 540  
 35 Pro Val Gln Tyr Pro Ala Asp Asp Val Thr Ala Glu Trp Gln Thr Thr  
 545 550 555 560  
 40 His Phe Asn Phe His Asp Ile Asp Glu Asn Val Leu Lys Leu Asp Val  
 565 570 575  
 Leu Gly His Asp Asp Pro Thr Met Ile Arg Lys Leu Gln Asp Leu Ser  
 580 585 590  
 Gly Ile Asp Pro Asn Lys Ile Pro Met Asp Asp Glu Gly Val Met Ala  
 595 600 605  
 50 Leu Phe Ser Gly Thr Asp Val Leu Gly Val Thr Pro Glu Gln Ile Gly  
 610 615 620  
 Thr Leu Arg Val Cys  
 625  
 55

## (2) INFORMATION FOR SEQ ID NO:222:

## (i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 693 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant

-312-

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

15 Met Ala Arg Glu Phe Ser Leu Glu Lys Thr Arg Asn Ile Gly Ile Met  
 1 5 10 15  
 Ala His Val Asp Ala Gly Lys Thr Thr Thr Thr Glu Arg Ile Leu Tyr  
 20 20 25 30  
 Tyr Thr Gly Lys Ile His Lys Ile Gly Glu Thr His Glu Gly Ala Ser  
 35 40 45  
 Gln Met Asp Trp Met Glu Gln Glu Gln Glu Arg Gly Ile Thr Ile Thr  
 50 55 60  
 25 Ser Ala Ala Thr Thr Ala Gln Trp Asn Asn His Arg Val Asn Ile Ile  
 65 70 75 80  
 Asp Thr Pro Gly His Val Asp Phe Thr Ile Glu Val Gln Arg Ser Leu  
 85 90 95  
 Arg Val Leu Asp Gly Ala Val Thr Val Leu Asp Ser Gln Ser Gly Val  
 100 105 110  
 35 Glu Pro Gln Thr Glu Thr Val Trp Arg Gln Ala Thr Glu Tyr Gly Val  
 115 120 125  
 Pro Arg Ile Val Phe Ala Asn Lys Met Asp Lys Ile Gly Ala Asp Phe  
 130 135 140  
 40 Leu Tyr Ser Val Ser Thr Leu His Asp Arg Leu Gln Ala Asn Ala His  
 145 150 155 160  
 Pro Ile Gln Leu Pro Ile Gly Ser Glu Asp Asp Phe Arg Gly Ile Ile  
 165 170 175  
 45 Asp Leu Ile Lys Met Lys Ala Glu Ile Tyr Thr Asn Asp Leu Gly Thr  
 180 185 190  
 50 Asp Ile Leu Glu Glu Asp Ile Pro Ala Glu Tyr Leu Asp Gln Ala Gln  
 195 200 205  
 Glu Tyr Arg Glu Lys Leu Ile Glu Ala Val Ala Glu Thr Asp Glu Glu  
 210 215 220  
 55 Leu Met Met Lys Tyr Leu Glu Gly Glu Glu Ile Thr Asn Glu Glu Leu  
 225 230 235 240  
 60 Lys Ala Gly Ile Arg Lys Ala Thr Ile Asn Val Glu Phe Phe Pro Val  
 245 250 255

-313-

	Leu	Cys	Gly	Ser	Ala	Phe	Lys	Asn	Lys	Gly	Val	Gln	Leu	Met	Leu	Asp	
				260					265					270			
5	Ala	Val	Ile	Asp	Tyr	Leu	Pro	Ser	Pro	Leu	Asp	Ile	Pro	Ala	Ile	Lys	
			275					280					285				
	Gly	Ile	Asn	Pro	Asp	Thr	Asp	Ala	Glu	Glu	Ile	Arg	Pro	Ala	Ser	Asp	
		290					295					300					
10	Glu	Glu	Pro	Phe	Ala	Ala	Leu	Ala	Phe	Lys	Ile	Met	Thr	Asp	Pro	Phe	
	305					310					315					320	
	Val	Gly	Arg	Leu	Thr	Phe	Phe	Arg	Val	Tyr	Ser	Gly	Val	Leu	Gln	Ser	
15					325					330					335		
	Gly	Ser	Tyr	Val	Leu	Asn	Thr	Ser	Lys	Gly	Lys	Arg	Glu	Arg	Ile	Gly	
				340					345					350			
20	Arg	Ile	Leu	Gln	Met	His	Ala	Asn	Ser	Arg	Gln	Glu	Ile	Asp	Thr	Val	
			355					360					365				
	Tyr	Ser	Gly	Asp	Ile	Ala	Ala	Ala	Val	Gly	Leu	Lys	Asp	Thr	Thr	Thr	
		370					375					380					
25	Gly	Asp	Ser	Leu	Thr	Asp	Glu	Lys	Ala	Lys	Ile	Ile	Leu	Glu	Ser	Ile	
	385					390					395					400	
	Asn	Val	Pro	Glu	Pro	Val	Ile	Gln	Leu	Met	Val	Glu	Pro	Lys	Ser	Lys	
30					405					410					415		
	Ala	Asp	Gln	Asp	Lys	Met	Gly	Ile	Ala	Leu	Gln	Lys	Leu	Ala	Glu	Glu	
				420					425					430			
35	Asp	Pro	Thr	Phe	Arg	Val	Glu	Thr	Asn	Val	Glu	Thr	Gly	Glu	Thr	Val	
			435					440					445				
	Ile	Ser	Gly	Met	Gly	Glu	Leu	His	Leu	Asp	Val	Leu	Val	Asp	Arg	Met	
		450					455				460						
40	Arg	Arg	Glu	Phe	Lys	Val	Glu	Ala	Asn	Val	Gly	Ala	Pro	Gln	Val	Ser	
	465					470					475					480	
	Tyr	Arg	Glu	Thr	Phe	Arg	Ala	Ser	Thr	Gln	Ala	Arg	Gly	Phe	Phe	Lys	
45					485					490				495			
	Arg	Gln	Ser	Gly	Gly	Lys	Gly	Gln	Phe	Gly	Asp	Val	Trp	Ile	Glu	Phe	
				500					505					510			
50	Thr	Pro	Asn	Glu	Glu	Gly	Lys	Gly	Phe	Glu	Phe	Glu	Asn	Ala	Ile	Val	
			515					520					525				
	Gly	Gly	Val	Val	Pro	Arg	Glu	Phe	Ile	Pro	Ala	Val	Glu	Lys	Gly	Leu	
		530					535					540					
55	Val	Glu	Ser	Met	Ala	Asn	Gly	Val	Leu	Ala	Gly	Tyr	Pro	Met	Val	Asp	
	545					550					555					560	
	Val	Lys	Ala	Lys	Leu	Tyr	Asp	Gly	Ser	Tyr	His	Asp	Val	Asp	Ser	Ser	
60					565					570					575		
	Glu	Thr	Ala	Phe	Lys	Ile	Ala	Ala	Ser	Leu	Ser	Leu	Lys	Glu	Ala	Ala	

-314-

580 585 590

Lys Ser Ala Gln Pro Ala Ile Leu Glu Pro Met Met Leu Val Thr Ile  
595 600 605

Thr Val Pro Glu Glu Asn Leu Gly Asp Val Met Gly His Val Thr Ala  
610 615 620

Arg Arg Gly Arg Val Asp Gly Met Glu Ala His Gly Asn Ser Gln Ile  
625 630 635 640

Val Arg Ala Tyr Val Pro Leu Ala Glu Met Phe Gly Tyr Ala Thr Val  
645 650 655

Leu Arg Ser Ala Ser Gln Gly Arg Gly Thr Phe Met Met Val Phe Asp  
660 665 670

His Tyr Glu Asp Val Pro Lys Ser Val Gln Glu Glu Ile Ile Lys Lys  
675 680 685

Asn Lys Gly Glu Asp  
690

(2) INFORMATION FOR SEQ ID NO:223:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 274 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

Ala Tyr Lys Gly His Gln Glu Tyr Val Leu Pro Gln Ala Ala Arg Lys  
1 5 10 15

Ile Tyr Ala Tyr Arg Arg Tyr Asp Leu Asn Glu Ser Pro Lys Thr Ala  
20 25 30

Leu Asp Leu Ile Ile Pro Asp Leu Phe Leu His Ile Leu Asn Pro Ala  
35 40 45

Glu Arg Glu Arg Lys Leu Lys Arg Glu Gly Val Glu Glu Leu Tyr Leu  
50 55 60

Leu Asp Phe Ser Ser Gln Phe Ala Ser Leu Thr Ala Gln Glu Phe Phe  
65 70 75 80

Ala Thr Tyr Ile Lys Ala Met Asn Ala Lys Ile Ile Val Ala Gly Phe  
85 90 95

Asp Tyr Thr Phe Gly Ser Asp Lys Lys Thr Ala Glu Asp Leu Lys Asp  
100 105 110

-315-

Tyr Phe Asp Gly Glu Val Ile Ile Val Pro Pro Val Glu Asp Glu Lys  
 115 120 125  
 5 Gly Lys Ile Ser Ser Thr Arg Ile Arg Gln Ala Ile Leu Asp Gly Asn  
 130 135 140  
 Val Lys Glu Ala Gly Lys Leu Leu Gly Ala Pro Leu Pro Ser Arg Gly  
 145 150 155 160  
 10 Met Val Val His Gly Asn Ala Arg Gly Arg Thr Ile Gly Tyr Pro Thr  
 165 170 175  
 Ala Asn Leu Val Leu Leu Asp Arg Thr Tyr Met Pro Ala Asp Gly Val  
 180 185 190  
 15 Tyr Val Val Asp Val Glu Ile Gln Arg Gln Lys Tyr Arg Ala Met Ala  
 195 200 205  
 Ser Val Gly Lys Asn Val Thr Phe Asp Gly Glu Glu Ala Arg Phe Glu  
 210 215 220  
 Val Asn Ile Phe Asp Phe Asn Gln Asp Ile Tyr Gly Glu Thr Val Met  
 225 230 235 240  
 25 Val Tyr Trp Leu Asp Arg Ile Arg Asp Met Thr Lys Phe Asp Ser Val  
 245 250 255  
 Asp Gln Leu Val Asp Gln Leu Lys Ala Asp Glu Glu Val Thr Arg Asn  
 260 265 270  
 30 Trp Ser

## (2) INFORMATION FOR SEQ ID NO:224:

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 124 amino acids  
 (B) TYPE: amino acid  
 40 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

50 Leu Arg Lys Glu Pro Ser Met Ala Lys Gly Glu Gly Lys Val Val Ala  
 1 5 10 15  
 55 Gln Asn Lys Lys Ala Arg His Asp Tyr Thr Ile Val Asp Thr Leu Glu  
 20 25 30  
 Ala Gly Met Val Leu Thr Gly Thr Glu Ile Lys Ser Val Arg Ala Ala  
 35 40 45  
 60 Arg Ile Asn Leu Lys Asp Gly Phe Ala Gln Val Lys Asn Gly Glu Val  
 50 55 60



-316-

Trp Leu Ser Asn Val His Ile Ala Pro Tyr Glu Glu Gly Asn Ile Trp  
 65 70 75 80

5 Asn Gln Glu Pro Glu Arg Arg Arg Lys Leu Leu Leu His Lys Lys Gln  
 85 90 95

Ile Gln Lys Leu Glu Gln Glu Thr Lys Gly Thr Gly Met Thr Leu Val  
 100 105 110

10 Pro Leu Lys Val Tyr Met Ala Thr Leu Ser Phe Phe  
 115 120

## (2) INFORMATION FOR SEQ ID NO:225:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 441 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 20 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

Ile Val Lys Glu Glu Lys Gly Leu Lys Glu Lys Gln Phe Trp Asn Arg  
 1 5 10 15

35 Ile Leu Glu Phe Ala Gln Glu Arg Leu Thr Arg Ser Met Tyr Asp Phe  
 20 25 30

40 Tyr Ala Ile Gln Ala Glu Leu Ile Lys Val Glu Glu Asn Val Ala Thr  
 35 40 45

Ile Phe Leu Pro Arg Ser Glu Met Glu Met Val Trp Glu Lys Gln Leu  
 50 55 60

45 Lys Asp Ile Ile Val Val Ala Gly Phe Glu Ile Tyr Asp Ala Glu Ile  
 65 70 75 80

Thr Pro His Tyr Ile Phe Thr Lys Pro Gln Asp Thr Thr Ser Ser Gln  
 85 90 95

50 Val Glu Glu Ala Thr Asn Leu Thr Leu Tyr Asp Tyr Ser Pro Lys Leu  
 100 105 110

55 Val Ser Ile Pro Tyr Ser Asp Thr Gly Leu Lys Glu Lys Tyr Thr Phe  
 115 120 125

Asp Asn Phe Ile Gln Gly Asp Gly Asn Val Trp Ala Val Ser Ala Ala  
 130 135 140

60 Leu Ala Val Ser Glu Asp Leu Ala Leu Thr Tyr Asn Pro Leu Phe Ile  
 145 150 155 160

-317-

Tyr Gly Gly Pro Gly Leu Gly Lys Thr His Leu Leu Asn Ala Ile Gly  
 165 170 175  
 5 Asn Glu Ile Leu Lys Asn Ile Pro Asn Ala Arg Val Lys Tyr Ile Pro  
 180 185 190  
 Ala Glu Ser Phe Ile Asn Asp Phe Leu Asp His Leu Arg Leu Gly Glu  
 195 200 205  
 10 Met Glu Lys Phe Lys Lys Thr Tyr Arg Ser Leu Asp Leu Leu Leu Ile  
 210 215 220  
 Asp Asp Ile Gln Ser Leu Ser Gly Lys Lys Val Ala Thr Gln Glu Glu  
 225 230 235 240  
 Phe Phe Asn Thr Phe Asn Ala Leu His Asp Lys Gln Lys Gln Ile Val  
 245 250 255  
 20 Leu Thr Ser Asp Arg Ser Pro Lys His Leu Glu Gly Leu Glu Glu Arg  
 260 265 270  
 Leu Val Thr Arg Phe Ser Trp Gly Leu Thr Gln Thr Ile Thr Pro Pro  
 275 280 285  
 25 Asp Phe Glu Thr Arg Ile Ala Ile Leu Gln Ser Lys Thr Glu His Leu  
 290 295 300  
 Gly Tyr Asn Phe Gln Ser Asp Thr Leu Glu Tyr Leu Ala Gly Gln Phe  
 305 310 315 320  
 Asp Ser Asn Val Arg Asp Leu Glu Gly Ala Ile Asn Asp Ile Thr Leu  
 325 330 335  
 35 Ile Ala Arg Val Lys Lys Ile Lys Asp Ile Thr Ile Asp Ile Ala Ala  
 340 345 350  
 Glu Ala Ile Arg Ala Arg Lys Gln Asp Val Ser Gln Met Leu Val Ile  
 355 360 365  
 40 Pro Ile Asp Lys Ile Gln Thr Glu Val Gly Asn Phe Tyr Gly Val Ser  
 370 375 380  
 Ile Lys Glu Met Lys Gly Ser Arg Arg Leu Gln Asn Ile Val Leu Ala  
 385 390 395 400  
 Arg Gln Val Ala Met Tyr Leu Ser Arg Glu Leu Thr Asp Asn Ser Leu  
 405 410 415  
 50 Pro Lys Ile Gly Lys Glu Leu Gly Glu Lys Ser Tyr His Ser His Ser  
 420 425 430  
 Cys Pro Cys Gln Asn Lys Ile Leu Asn  
 435 440  
 55

(2) INFORMATION FOR SEQ ID NO:226:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 201 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

60

-318-

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:

15 Glu Leu Val Ser Thr Met Tyr Phe Arg Phe Asp Tyr Tyr Ser Gln Asn  
 1 5 10 15  
 Leu Gly Glu Ile Phe Ala Ile Gly Met Val Val Gly His Leu Arg Trp  
 20 20 25 30  
 Leu Ile Thr Gly Ala Leu Val Leu Tyr Ile Phe Ala Asp Arg Lys Leu  
 35 40 45  
 Ile Asn Thr Trp Asp Phe Leu Asp Ile Ala Ala Pro Ser Val Met Ile  
 50 55 60  
 25 Ala Gln Ser Leu Gly Arg Trp Gly Asn Phe Phe Asn Gln Glu Ala Tyr  
 65 70 75 80  
 Gly Ala Thr Val Asp Asn Leu Asp Tyr Leu Pro Gly Phe Ile Arg Asp  
 85 90 95  
 Gln Met Tyr Ile Glu Gly Ser Tyr Arg Gln Pro Thr Phe Leu Tyr Glu  
 100 105 110  
 35 Ser Leu Trp Asn Leu Leu Gly Phe Ala Leu Ile Leu Ile Phe Arg Arg  
 115 120 125  
 Lys Trp Lys Ser Leu Arg Arg Gly His Ile Thr Ala Phe Tyr Leu Ile  
 130 135 140  
 40 Trp Tyr Gly Phe Gly Arg Met Val Ile Glu Gly Met Arg Thr Asp Ser  
 145 150 155 160  
 Leu Met Phe Phe Gly Leu Arg Val Ser Gln Trp Leu Ser Val Val Leu  
 165 170 175  
 45 Ile Gly Leu Gly Ile Met Ile Val Ile Tyr Gln Asn Arg Lys Lys Ala  
 180 185 190  
 50 Pro Tyr Tyr Ile Thr Glu Glu Glu Asn  
 195 200

(2) INFORMATION FOR SEQ ID NO:227:

55 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 491 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

60

(ii) MOLECULE TYPE: peptide

-319-

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:

10 Leu Glu Asp Phe Pro Leu Ser Val Thr Asn Pro Tyr Gly Arg Thr Lys  
 1 5 10 15  
 Leu Met Leu Glu Glu Ile Leu Thr Asp Ile Tyr Lys Ala Asp Ser Glu  
 20 25 30  
 15 Trp Asn Val Val Leu Leu Arg Tyr Phe Asn Pro Ile Gly Val His Glu  
 35 40 45  
 Ser Gly Asp Leu Gly Glu Asn Pro Asn Gly Ile Pro Asn Asn Leu Leu  
 20 50 55 60  
 Pro Tyr Val Thr Gln Val Ala Val Gly Lys Leu Glu Gln Val Gln Val  
 65 70 75 80  
 25 Phe Gly Asp Asp Tyr Asp Thr Glu Asp Gly Thr Gly Val Arg Asp Tyr  
 85 90 95  
 Ile His Val Val Asp Leu Ala Lys Gly His Val Ala Ala Leu Lys Lys  
 100 105 110  
 30 Ile Gln Lys Gly Ser Gly Leu Asn Val Tyr Asn Leu Gly Thr Gly Lys  
 115 120 125  
 Gly Tyr Ser Val Leu Glu Ile Ile Gln Asn Met Glu Lys Ala Val Gly  
 130 135 140  
 35 Cys Pro Ile Pro Tyr Arg Ile Val Glu Arg Arg Pro Gly Asp Ile Ala  
 145 150 155 160  
 40 Ala Cys Tyr Ser Asp Pro Ala Lys Ala Lys Ala Glu Leu Gly Trp Glu  
 165 170 175  
 Ala Glu Leu Asp Ile Thr Gln Met Cys Glu Gly His Gly Val Gly Arg  
 180 185 190  
 45 Ala Ser Ile Gln Met Asp Leu Lys Thr Lys Met Met Ile Ser Ile Ile  
 195 200 205  
 Val Pro Cys Leu Asn Glu Glu Glu Val Leu Pro Leu Phe Tyr Gln Ala  
 210 215 220  
 50 Leu Glu Ala Leu Leu Pro Asp Leu Glu Thr Glu Ile Glu Tyr Val Phe  
 225 230 235 240  
 55 Val Asp Asp Gly Ser Ser Asp Gly Thr Leu Glu Leu Leu Lys Ala Tyr  
 245 250 255  
 Arg Glu Gln Asn Pro Ala Val His Tyr Ile Ser Phe Ser Arg Asn Phe  
 260 265 270  
 60 Gly Lys Glu Ala Ala Leu Tyr Ala Gly Leu Gln Tyr Ala Thr Gly Asp  
 275 280 285

-320-

Leu Val Val Val Met Asp Ala Asp Leu Gln Asp Pro Pro Ser Met Leu  
 290 295 300  
 5 Phe Glu Met Lys Asn Val Leu Asp Lys Asn Val Asp Leu Asp Cys Val  
 305 310 315 320  
 Gly Thr Arg Arg Thr Ser Arg Glu Gly Glu Pro Phe Phe Arg Ser Phe  
 325 330 335  
 10 Cys Ala Val Leu Phe Tyr Arg Leu Met Gln Lys Ile Ser Pro Val Ala  
 340 345 350  
 15 Leu Pro Ser Gly Val Arg Asp Phe Arg Met Met Arg Arg Ser Val Val  
 355 360 365  
 Asp Ala Ile Leu Ser Leu Thr Glu Ser Asn Arg Phe Ser Lys Gly Leu  
 370 375 380  
 20 Phe Ala Trp Val Gly Phe Lys Thr His Tyr Leu Asp Tyr Pro Asn Val  
 385 390 395 400  
 Glu Arg Gln Ala Gly Lys Thr Ser Trp Ser Phe Arg Gln Leu Phe Phe  
 405 410 415  
 25 Tyr Ser Ile Glu Gly Ile Val Asn Phe Ser Asp Phe Pro Leu Thr Ile  
 420 425 430  
 30 Ala Phe Val Ala Gly Leu Leu Ser Cys Phe Leu Ser Leu Leu Met Thr  
 435 440 445  
 Phe Phe Val Val Val Arg Thr Leu Ile Leu Gly Asn Pro Thr Ser Gly  
 450 455 460  
 35 Trp Thr Ser Leu Met Ala Val Ile Leu Tyr Leu Gly Gly Ile Gln Leu  
 465 470 475 480  
 Leu Thr Ile Gly Ile Leu Gly Lys Tyr Asn Gln  
 485 490

## (2) INFORMATION FOR SEQ ID NO:228:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 277 amino acids  
 45 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide  
 50

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:228:

Val Ile Ile Ile Asp Asp Asn Tyr Ser Asn Val Asn Leu Arg Asn Lys  
 1 5 10 15

60

Ile Ile His Gln Phe Gly Tyr Thr Asn His Arg Ile Lys Leu Ile Leu

-321-

	20	25	30
5	Ser Asn Glu Asp Leu Gly Ala Thr Asn Ala Arg Asn Ile Gly Ile Lys 35 40 45		
	Asn Ser Arg Gly Lys Tyr Ile Ser Phe Leu Asp Asp Asp Asp Glu Tyr 50 55 60		
10	Met Pro Asp Arg Ile Leu Lys Leu Met Ala Cys Phe Lys Lys Ser Arg 65 70 75 80		
	Met Lys Asn Leu Ala Leu Val Tyr Ser Tyr Gly Ile Ile Ile Tyr Pro 85 90 95		
15	Asn Gly Thr Arg Glu Glu Glu Lys Thr Asp Phe Val Gly Asn Pro Leu 100 105 110		
	Phe Val Gln Met Val His Asn Ile Ala Gly Thr Ser Phe Trp Leu Cys 115 120 125		
20	Lys Lys Glu Val Leu Glu Leu Ile Asn Gly Phe Glu Lys Ile Asp Ser 130 135 140		
	His Gln Asp Gly Val Val Leu Leu Lys Leu Leu Ala Gln Gly Tyr Gln 145 150 155 160		
25	Ile Asp Ile Val Arg Glu Phe Leu Val Asn Tyr Tyr Ala His Ser Lys 165 170 175		
	Glu Asn Gly Ile Thr Gly Val Thr Gln Lys Thr Ile Asn Ala Asp Glu 180 185 190		
30	Glu Tyr Tyr Asn Tyr Cys Arg Lys Tyr Phe Asn Leu Leu Ser Phe Asn 195 200 205		
35	Glu Arg Ile Leu Val Thr Lys Lys Tyr Tyr Ser Leu Asn Ile Lys Arg 210 215 220		
	Leu Leu Leu Ile Gly Asp Lys Cys Lys Ala Leu Lys Val Ile Lys Lys 225 230 235 240		
40	Ala Arg Glu Glu Lys Ile Phe Asn Glu Phe Leu Phe Leu Lys Tyr Met 245 250 255		
	Leu Leu Tyr Arg Ser Phe Phe Tyr Cys Ile Tyr Asp Asn Tyr Val Gln 260 265 270		
45	Leu Lys Phe Arg Lys 275		

-322-

## CLAIMS

1. An isolated nucleic acid compound comprising a sequence identical to or substantially identical to a  
5 sequence selected from the group consisting of SEQ ID NO:1 through SEQ ID NO:86.

2. An isolated nucleic acid compound comprising a sequence identical to or substantially identical to a  
10 sequence selected from the group consisting of SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, and SEQ  
15 ID NO:121.

3. A substantially purified protein or fragment thereof from *S. pneumoniae* wherein said protein is selected from the group consisting of SEQ ID NO:88, SEQ ID NO:90, SEQ  
20 ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, and SEQ ID NO:123 through SEQ ID NO:228.

25

4. An isolated nucleic acid compound encoding any one of the proteins or fragments thereof of Claim 3.

5. A vector comprising any one of the nucleic acid  
30 compounds of claims 1, 2, or 4.

6. A recombinant host containing any one of the vectors of claim 5.

-323-

7. A substantially purified protein from *Streptococcus pneumoniae* as in Claim 3 wherein said protein is an external target protein selected from Table 1.

5           8. A substantially purified protein from *Streptococcus pneumoniae* as in Claim 3 wherein said protein is a hypothetical protein selected from Table 1.

9. A substantially purified protein from *Streptococcus pneumoniae* as in Claim 3 wherein said protein is a cell wall synthetic protein selected from Table 1.

10           10. A substantially purified protein from *Streptococcus pneumoniae* as in Claim 3 wherein said protein is a minimal gene set protein selected from Table 1.

11. A DNA chip having arrayed thereon any at least 15 base pair fragment of any one or more of the nucleic acids of claim 1.

20           12. A DNA chip having arrayed thereon any at least 15 base pair fragment of any one or more of the nucleic acids of claim 2.

25           13. A method for evaluating gene expression in *Streptococcus pneumoniae* comprising the step of incubating a DNA chip of claim 11 or Claim 12 with cDNA prepared from *Streptococcus pneumoniae* under conditions suitable for hybridization of complementary nucleic acid sequences.

30           14. A computer readable medium having recorded thereon any one or more of the nucleotide sequences of Claims 1 or Claim 2.



-324-

15. A method for identifying virulence genes in *S. pneumoniae*, comprising the steps of:

- a) preparing a DNA chip as in claim 11,
- b) preparing labeled cDNAs from
  - 5           i) *S. pneumoniae* cells recovered from an in vivo environment, and
  - ii) *S. pneumoniae* cells grown in vitro,
- c) hybridizing individually the cDNAs of steps (b) (i) and (b) (ii) to a chip of step (a); and
- 10           d) identifying a genomic DNA fragment or fragments on said chip that hybridize to the cDNAs of (b) (i) but not with the cDNAs of (b) (ii).

16. An antibody that selectively binds to a  
15 protein or peptide of Claim 3.

17. An antibody that selectively binds to an external target protein, or fragment thereof, identified in Table 1.

20           18. A DNA chip of Claim 11 or Claim 12 further comprising a layer of *S. pneumoniae* cells wherein said layer contacts with said nucleic acids.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/22578

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : 536/23.7, 23.1; 530/350, 387.1; 435/320.1, 252.3, 6; 360/1

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.7, 23.1; 530/350, 387.1; 435/320.1, 252.3, 6; 360/1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN, APS, chip#, dna#, oligo?

MPSRCH of nucleic acid and amino acid sequence databases of IntelliGenetics (SEQ ID NOs 1-228)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	FUKUNAGA et al. Expression cloning of a receptor for murine granulocyte colony-stimulating factor. Cell. 20 April 1990, Vol. 61, pages 341-350, see entire document.	11-13 and 15
Y	GEETHA-HABIB et al. Glycosylation site binding protein, a component of oligosaccharyl transferase, is highly similar to three other 57 kd liminal proteins of the ER. Cell. 23 September 1988, Vol. 54, pages 1053-1060, see entire document.	11-13 and 15
Y	STRAUSS et al. Complete nucleotide sequence of the genomic RNA of Sindbis virus. Virology. February 1984, Vol. 133, No. 1, pages 92-110, see entire document.	11-13 and 15

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	* T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
* A* document defining the general state of the art which is not considered to be of particular relevance	* X	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
* B* earlier document published on or after the international filing date	* Y	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
* L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	* A*	document member of the same patent family
* O* document referring to an oral disclosure, use, exhibition or other means		
* P* document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

25 MARCH 1998

Date of mailing of the international search report

09 APR 1998

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JAMES MARTINELL

Telephone No. (703) 308-8196

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/22578

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WILSON et al. 2.2 Mb of contiguous nucleotide sequence from chromosome III of <i>C. elegans</i> . <i>Nature</i> . 03 March 1994, Vol. 368, No. 6466, pages 32-38, see entire document.	11-13 and 15
Y	RESTREPO et al. Antigen diversity in the bacterium <i>B. hermsii</i> through "somatic" mutations in rearranged <i>vmp</i> genes. <i>Cell</i> . 09 September 1994, Vol. 78, pages 867-876, see entire document.	11-13 and 15
Y	WHITEWAY et al. Dominant negative selection of heterologous genes: Isolation of <i>Candida albicans</i> genes that interfere with <i>Saccharomyces cerevisiae</i> mating factor-induced cell cycle arrest. <i>Proc. Natl. Acad. Sci. USA</i> . 15 October 1992, Vol 89, No. 20, pages 9410-9414, see entire document.	11-13 and 15
Y	LACKS et al. Genetic basis of the complementary DpnI and DpnII restriction systems of <i>S. pneumoniae</i> : An intracellular cassette mechanism. <i>Cell</i> . 26 September 1986, Vol. 46, pages 993-1000, see entire document.	11-13 and 15
Y	STEENBERGH et al. Structure and expression of the human calcitonin/CGRP genes. <i>FEBS Letters</i> . 01 December 1986, Vol. 209, No. 1, pages 97-103, see entire document.	11-13 and 15
Y	SHAMBAUGH et al. The spliceosomal U small nuclear RNAs of <i>Ascaris lumbricoides</i> . <i>Molecular and Biochemical Parasitology</i> . April 1994, Vol. 64, No. 2, pages 349-352, see entire document.	11-13 and 15
Y	MAXWELL et al. A novel NF-kappaB p65 spliced transcript lacking exons 6 and 7 in a non-small cell lung carcinoma cell line. <i>Gene</i> . 12 December 1995, Vol. 166, No. 2, pages 339-340, see entire document.	11-13 and 15
Y	KONYECSNI et al. Broad-host-range plasmid and M13 bacteriophage-derived vectors for promoter analysis in <i>Escherichia coli</i> and <i>Pseudomonas aeruginosa</i> . <i>Gene</i> . 30 December 1988, Vol. 74, No. 2, pages 375-386, see entire document.	11-13 and 15
Y	ALLIKMETS et al. Cloning and organization of the <i>abc</i> and <i>mdl</i> genes of <i>Escherichia coli</i> ; Relationship to eukaryotic multidrug resistance. <i>Gene</i> . 22 December 1993, Vol. 136, Nos. 1 and 2, pages 231-236, see entire document.	11-13 and 15
Y	SCHWEIZER. Two plasmids, X1918 and Z1918, for easy recovery of the <i>xylE</i> and <i>lacZ</i> reporter genes. <i>Gene</i> . 30 November 1993, Vol. 134, No. 1, pages 89-91, see entire document.	11-13 and 15

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/22578

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KRAFT et al. Sequence of the complete P protein gene and part of the M protein gene form the histidine transport operon of <i>Escherichia coli</i> compared to that of <i>Salmonella typhimurium</i> . <i>Nucleic Acids Res.</i> 26 October 1987, Vol. 15, No. 20, page 8568, see entire document.	11-13 and 15
Y	SCHULER et al., Characterization of the human gene encoding LBR, an integral protein of the nuclear envelope inner membrane., <i>J. Biol. Chem.</i> 15 April 1994, Vol. 269, No. 15, pages 11312-11317, see entire document.	11-13 and 15
Y	LOTTENBERG et al. Cloning, sequence analysis, and expression in <i>Escherichia coli</i> of a streptococcal plasmin receptor. <i>J. Bacteriol.</i> August 1992, Vol. 174, No. 16, pages 5204-5210, see entire document.	11-13 and 15
Y	COPPOLA et al. Sequence and transcriptional activity of the <i>Escherichia coli</i> K-12 chromosome region between <i>rrnC</i> and <i>ilvGMEDA</i> . <i>Gene.</i> 02 January 1991, Vol. 97, No. 1, pages 21-27, see entire document.	11-13 and 15
Y	MUTO et al. Preferential use of A- and U-rich codons for <i>Mycoplasma capricolum</i> ribosomal proteins S8 and L6. <i>Nucleic Acids Res.</i> 1984, Vol. 12, No. 12, pages 8209-8217, see entire document.	11-13 and 15
Y	RUDNER et al. The <i>spo0K</i> locus of <i>Bacillus subtilis</i> is homologous to the oligopeptide permease locus is required for sporulation and competence. <i>J. Bacteriology.</i> February 1991, Vol. 173, No. 3, pages 1388-1398, see entire document.	11-13 and 15
Y	GREEN et al. Isolation and characterisation of a bovine cDNA encoding eukaryotic initiation factor 2 $\alpha$ . <i>Biochim. Biophys. Acta.</i> 08 October 1991, Vol. 1090, No. 2, pages 277-280, see entire document.	11-13 and 15
Y	LOGEMANN et al. Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. <i>Proc. Natl. acad. Sci. USA.</i> 20 June 1995, Vol. 92, No. 13, pages 5905-5909, see entire document.	11-13 and 15
Y	SHERWOOD et al. Characterization of HIR1 and HIR2, two genes required for regulation of histone gene transcription in <i>Saccharomyces cerevisiae</i> . <i>Molec. Cell. Biol.</i> January 1993, Vol. 13, No. 1, pages 28-38, see entire document.	11-13 and 15

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/22578

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GAGNON et al. Cloning, sequencing and expression in <i>Escherichia coli</i> of the <i>ptsI</i> gene encoding enzyme I of the phosphoenolpyruvate:sugar phosphotransferase transport system from <i>Streptococcus salivarius</i> . <i>Gene</i> . 02 November 1992, Vol. 121, No. 1, pages 71-78, see entire document.	11-13 and 15-17
Y	PUYET et al. Characterization of the <i>Streptococcus pneumoniae</i> maltosaccharide regulator MalR, a member of the LacI-GalR family of repressors displaying distinctive genetic features. <i>J. Biol. Chem.</i> 05 December 1993, Vol. 268, No. 34, pages 25402-25408, see entire document.	11-13 and 15
Y	CROSS et al. Purification of CpG islands using a methylated DNA binding column. <i>Nature Genetics</i> . March 1994, Vol. 6, pages 236-244, see entire document.	11-13 and 15
Y	SALUJA et al. The genetic basis of colony opacity in <i>Streptococcus pneumoniae</i> : evidence for the effect of box elements on the frequency of phenotypic variation. <i>Molecular Microbiology</i> . April 1995, Vol. 16, No. 2, pages 215-227, see entire document.	11-13 and 15-17
Y	ROTHNIE et al. Retrotransposon-like nature of Tpl elements: implications for the organisation of highly repetitive, hypermethylated DNA in the genome of <i>Physarum polycephalum</i> . <i>Nucleic Acids Res.</i> 25 January 1991, Vol. 19, No. 2, pages 279-286, see entire document.	11-13 and 15
Y	GITT et al. A strong sequence homology exists between the major RNA polymerase $\delta$ factors of <i>Bacillus subtilis</i> and <i>Escherichia coli</i> . <i>J. Biol. Chem.</i> 25 June 1985, Vol. 260, No. 12, pages 7178-7185, see entire document.	11-13 and 15
Y	LEUVEN et al. Structure of the gene (LRP1) coding for the human $\alpha 2$ -macroglobulin receptor lipoprotein receptor-related protein. <i>Genomics</i> . 01 November 1994, Vol. 24, No. 1, pages 78-89, see entire document.	11-13 and 15
Y	YOSHIOKA et al. Nucleotide sequence of the promoter-distal region of the <i>tra</i> operon of plasmid R100, including <i>traI</i> (DNA helicase I) and <i>traD</i> genes. <i>J. Mol. Biol.</i> 05 July 1990, Vol. 214, No. 1, pages 39-53, see entire document.	11-13 and 15
Y	HUI et al. Genetic transformation in <i>Streptococcus pneumoniae</i> : Nucleotide sequence analysis shows <i>comA</i> , a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family. <i>J. Bacteriol.</i> January 1991, Vol. 173, No. 1, pages 372-381, see entire document.	11-13 and 15

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/22578

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BARON et al. Co-regulation of two gene activities by tetracycline via a bidirectional promoter. Nucleic Acids Res. 11 September 1995, Vol. 23, No. 17, pages 3605-3606, see entire document.	11-13 and 15
Y	BORK et al. Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiology. Molecular Microbiology. June 1995, Vol. 16, No. 5, pages 955-967, see entire document.	11-13 and 15
Y	CHAMBERS et al. The pMTL nic <sup>+</sup> cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. Gene. 15 August 1988, Vol. 68, No. 1, pages 139-149, see entire document.	11-13 and 15
Y	FENOLL et al. Molecular basis of the optochin-sensitive phenotype of pneumococcus: characterization of the genes encoding the F <sub>0</sub> complex of the Streptococcus pneumoniae and Streptococcus oralis H <sup>+</sup> -ATPases, Molecular Microbiology, May 1994, Vol. 12, No. 4, pages 587-598, see entire document.	11-13 and 15-17
Y	STEGLITZ-MORSDORF et al. Cloning, heterologous expression, and sequencing of the Proteus vulgaris glnAntrBC operon and implications of nitrogen control on heterologous urease expression. FEMS Microbiology Letters. 15 January 1993, Vol. 106, No. 2, pages 157-164, see entire document.	11-13 and 15
Y	DE WIT et al. Structure of the gene for the human myeloid IgA Fc receptor (CD89). J. Immunol. 01 August 1995, Vol. 155, No. 3, pages 1203-1209, see entire document.	11-13 and 15
Y	WEBSTER et al. Nucleotide sequence of the galactose gene cluster of Kluyveromyces lactis, Nucleic Acids Res. 25 August 1988, Vol. 16, No. 16, pages 8192-8194, see entire document.	11-13 and 15
Y	KIRSCH et al. Chemotactic methylesterase promotes adaptation to high concentrations of attractant in Bacillus subtilis. J. Biol. Chem. 05 September 1993, Vol. 268, No. 25, pages 18610-18616, see entire document.	11-13 and 15
Y	YAMASHITA et al. Molecular characterization of a Streptococcus mutans mutant altered in environmental stress responses. J. Bacteriol. October 1993, Vol. 175, No. 19, pages 6220-6228, see entire document.	11-13 and 15-17
Y	FRASER et al. The minimal gene complement of Mycoplasma genitalium. Science. 20 October 1995, Vol. 270, pages 397-403, see entire document.	11-13 and 15

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/22578

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	TOMINAGA et al. Upstream region of a genomic gene for human mitochondrial transcription factor 1. Biochim. Biophys. Acta. 15 June 1992, Vol. 1131, No. 2, pages 217-219, see entire document.	11-13 and 15
Y	CORTI et al. Molecular cloning of cDNAs encoding human carnitine acetyltransferase and mapping of the corresponding gene to chromosome 9q34.1. Genomics. 01 September 1994, Vol. 23, No. 1, pages 94-99, see entire document.	11-13 and 15
Y	LUSIS et al. Cloning and expression of apolipoprotein B, the major protein of low and very low density lipoproteins. Proc. Natl. Acad. Sci. USA. July 1985, Vol. 82, No. 14, pages 4597-4601, see entire document.	11-13 and 15
Y	KOHLBRECHER et al. Staphylococcal phosphoenolpyruvate-dependent phosphotransferase system: Molecular cloning and nucleotide sequence of the Staphylococcus carnosus ptsI gene and expression and complementation studies of the gene product. J. Bacteriol. April 1992, Vol. 174, No. 7, pages 2208-2214, see entire document.	16 and 17
Y	FLEISCHMANN et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science. 28 July 1995, Vol. 269, pages 496-512, see entire document.	11-13 and 15-17
Y	PEASE et al. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. Proc. Natl. Acad. Sci. USA. 24 May 1994, Vol. 91, No. 11, pages 5022-5026, see entire document.	11-13 and 15
X,P	CHENG et al. The com locus controls genetic transformation in Streptococcus pneumoniae. Molecular Microbiology. February 1997, Vol. 23, No. 4, pages 683-692, see entire document.	1, 4-6, and 14
Y	MORIARTY et al. Antibodies to peptides detect new hepatitis B antigen: Serological correlation with hepatocellular carcinoma. Science. 25 January 1985, Vol. 227, pages 429-433, see entire document.	16 and 17
X	WO 96/10647 A (FUSO PHARMACEUTICAL INDUSTRIES, LTD.) 11 April 1996, see entire document.	1, 4-6, and 14
X — Y	ALLOING et al. The ami locus of the Gram-positive bacterium Streptococcus pneumoniae is similar to binding protein-dependent transport operons of Gram-negative bacteria. Molecular Microbiology. 1990, Vol. 4, No. 4, pages 633-644, see entire document.	1, 4-6, and 14 ----- 16 and 17

**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/US97/22578

**A. CLASSIFICATION OF SUBJECT MATTER:**  
IPC (6):

C12N 15/31, 15/00, 1511; C07K 14/315; C07H 21/04; A61K 39/40, 39/395; G01N 33/569; G11B 5/00